

Delusions of Virtue: Kant on Self-Conceit

KATE MORAN

Brandeis University

Email: kmoran@brandeis.edu

Abstract

Little extended attention has been given to Kant's notion of self-conceit (*Eigendünkel*), though it appears throughout his theoretical and practical philosophy. Authors who discuss self-conceit often describe it as a kind of imperiousness or arrogance in which the conceited agent seeks to impose selfish principles upon others, or sees others as worthless. I argue that these features of self-conceit are symptoms of a deeper and more thoroughgoing failure. Self-conceit is best described as the tendency to insist upon one's own theoretical or practical conclusions at any cost, while still wanting to appear – to oneself or to others – as though one is abiding by the constraints of theoretical or practical reason. Self-conceit is thus less centrally the tendency to impose one's will or inclinations upon others, and more centrally the tendency to reconstruct evidence and rationalize so that one may be convinced of one's own virtue. While the conceited agent may ultimately impose her judgement upon others, she does so in order to preserve her delusion of virtue.

Keywords: Kant, Immanuel, self-conceit, self-love, self-deception, arrogance

1. Introduction

In chapter 3 of the *Analytic of the Critique of Practical Reason* (On the Incentives of Pure Practical Reason), Kant makes a distinction between two kinds of self-regard that an agent might have. The first of these is self-love (*Selbstliebe*, *Eigenliebe*); the other is self-conceit (*Eigendünkel*). Kant explains that self-love is characterized by the tendency to turn 'subjective determining grounds of choice into the objective determining ground of the will in general'.¹ Self-conceit, on the other hand, is characterized by the tendency to make self-love 'itself lawgiving and the unconditional practical principle'. The moral law 'infringes upon' self-love, but it 'strikes down' self-conceit (Kant 1996a: 199; *KpV*, 5: 73).

The distinction seems straightforward on its face, but questions soon arise. First, how shall we understand the distinction between the agent who quite acceptably treats certain inclinations as ‘subjective determining grounds of choice’ and the agent who treats inclination as the ‘*objective* determining ground of the will in general’? At first blush, the answer seems easy enough: the agent who treats inclination as ‘the objective determining ground of the will’ replaces the moral law with grounds provided by inclination. But this answer would seem to obscure the distinction between self-love and self-conceit sketched above. Self-*conceit*, according to that distinction, is the tendency to make self-love the ‘unconditional practical principle’. Kant appears to locate two distinct kinds of moral error in this passage, but it is not at all clear what these types of error are, or how they are meant to operate alongside each other.

These observations prompt a related question, namely, whether moral error can be a failure of self-love without also being a failure of self-conceit. Kant’s description of self-love in the second *Critique* makes reference to the agent’s reasons for action: an agent who evinces unreasonable self-love wrongly assesses her inclinations as objectively valid reasons for action. And while moral choices can be described this way, these decisions are equally well described in terms of the principles according to which we constrain ourselves and expect to constrain others. If I am running behind schedule and decide to drive my car in a lane reserved for bicycles, how should we describe my action? I am, after all, treating my desire to be on time as a decisive reason to drive in the bike lane. But roads are populated by other agents and governed by a system of rules. In this context, my act is equally an implicit statement about my attitude toward these agents and the practical principles that govern our interactions. If self-love has to do with *reasons* for action and self-conceit has to do with *principles* of action, we may legitimately wonder if moral failure ever occurs on one level, but not the other.

As far as Kant’s sketch in the *Critique of Practical Reason* goes, it is difficult to discern the distinction or the relationship between self-love and self-conceit. Fortunately, Kant’s discussions of self-conceit in other texts shed light on his understanding of the concept. Before analysing these texts, however, I survey the most common interpretation of moral self-conceit in the secondary literature. On this account, self-conceit differs from self-love in the conceited agent’s tendency to impose her principles and judgements upon others, or in her tendency to see others as worthless.²

2. Self-Conceit as Moral Imperiousness

Lewis White Beck, Andrews Reath and Stephen Darwall each emphasize Kant's claim that self-conceit turns self-love into 'the unconditional practical principle' and characterize self-conceit as a tendency to impose one's self-love upon others, or to regard oneself as uniquely and unilaterally legislating. Beck glosses self-conceit as 'the inclination to regard one's own subjective maxims and interests as having the authority of a law' (Beck 1960: 219). Though Beck does not explicitly mention any attitude toward others in his account of self-conceit, it is easy to imagine that a sense of imperiousness might develop from such an inclination. Reath makes this attitude towards others explicit when he describes self-conceit as 'a tendency to treat oneself or one's inclinations as providing reasons for the actions of *others*' (Reath 2006: 15). Reath explains that self-conceit is thus marked by a desire to 'dominate' others and 'treat [one's own] inclinations as special sources of reasons or value' (16–17). Similarly, Darwall observes that self-conceit is a tendency to think that one has a 'normative standing that others don't have to dictate reasons just because of who or what one is'. This is importantly different from the type of authority that 'an especially good advisor' might have. Rather, it is the 'fantasy that one has a fundamental "lawgiving" standing that others simply don't have – as if one were a king or God'.³

Allen Wood also argues that self-conceit is a 'propensity to regard our own inclinations as legislative' and concludes, further, that the conceited agent 'thereby [places his] comparative self-worth above the absolute worth other rational beings have according to the moral law' (Wood 1996: 154). Wood bases his analysis on an important observation: Kant describes self-love and self-conceit as subspecies of *Wohllollen* (benevolence, well-wishing) and *Wohlgefallen* (well-liking), respectively. Specifically, self-regard (*Selbstsucht*) manifests itself either as a 'pre-dominant benevolence toward oneself' (*Wohllollen gegen sich selbst*) or as a 'satisfaction with oneself' (*Wohlgefallen an sich selbst*) (Kant 1996a: 199; *KpV*, 5: 73). The former is self-love (*Selbstliebe*, *Eigenliebe*), while the latter is self-conceit (*Eigendiinkel*). Wood points out that in the later *Metaphysics of Morals*, Kant cites well-liking (*Wohlgefallen*) of others as the ground of other-directed benevolence or well-wishing (*Wohllollen*). Accordingly, Wood argues for an analogy between other-directed and self-directed attitudes. Just as *Wohlgefallen* is the ground of *Wohllollen* in the context of other-directed attitudes, so is self-oriented *Wohlgefallen* (self-conceit) the ground of self-oriented *Wohllollen* (self-love) in the case of selfish moral error (Wood 1996: 145, 148). Of course, as Wood correctly notes, particular instances of 'reasonable self-love' are

perfectly permissible (155). But the self-directed *Wohlgefallen* associated with self-conceit causes the agent to make self-love her fundamental maxim, or to make the fundamental choice to subordinate the moral law to self-love. This leads Wood to conclude that self-conceit *just is* radical evil and the ‘propensity to regard our own inclinations as legislative’ (154).

In a seeming reversal of Wood’s account, Stephen Engstrom argues that self-conceit *arises out of* self-love. Self-conceit, Engstrom argues, is ‘not so directly tied to our nature as self-love’, because Kantian self-love is a more primitive form of self-regard, while self-conceit seems to depend in crucial ways upon an awareness of the moral law (Engstrom 2010: 104). Engstrom argues that Kant’s description of self-conceit as a type of self-directed *Wohlgefallen* suggests that self-conceit should be understood in terms of esteem. Esteem, especially in the language of the third *Critique*, is an inherently comparative notion. As Engstrom puts it, ‘the self-esteem constituting self-conceit rests essentially on comparative aesthetic judgements asserting one’s superiority over others’ (109). Thus, self-conceit is a ‘representation of oneself [that] includes a comparison with others and in particular a placing of oneself above them’.⁴

These authors are, I think, correct in their observation that self-conceit involves a kind of imperiousness or arrogance. However, I argue that this feature of self-conceit is a symptom of a more fundamental failure of reasoning. Specifically, the conceited agent insists upon her own theoretical or practical conclusions at any cost, while still wanting to appear – to herself or to others – as though she is abiding by the constraints of theoretical or practical reason. In all cases of self-conceit, the conceited individual ignores a crucial point that Thomas Hill has observed about ‘standards of rationality’ – namely, ‘that practical reason, like theoretical reason, enables us to reach conclusions on some basis other than that they get us what we most want’.⁵ Lacking any principled way to apply the standards of theoretical and practical reason, the conceited agent constructs a chimera of objective validity that is nearly impossible to set right. Self-conceit is thus not the decision to *ignore* the constraint of impartiality and impose one’s own conclusions upon others. Rather, it is the tendency to construct a fiction that one is abiding by this constraint when one has actually failed to heed it. The failure of self-conceit is thus more intractable and insidious than the imposition of one’s will, or the claim to superiority, even if it often ultimately takes this form. It is no accident that Kant associates the failure with a kind of delusion (*Wahn, Wahnsinn*).⁶

The remainder of this discussion examines Kant's notion of self-conceit by investigating two key sources of information about the concept. Section 3 explores the habits and tendencies particular to self-conceit. Section 4 examines Kant's analyses of moral self-conceit. The concluding section returns to the puzzles raised in section 1.

3. Conceited Tendencies

Overstepping and fanaticism

One of the most prominent tendencies that Kant associates with self-conceit is the propensity to overstep or exceed (*übersteigen*) the boundaries of theoretical or practical reason.⁷ A nascent, and perhaps innocent, form of this error occurs within the context of pedagogy. Kant observes that educators are often forced to present knowledge to students that 'outstrips' students' maturity and ability. Such premature presentation of knowledge is the source of 'intractable' and 'absurd' prejudices in students, and equally 'the source of the precocious prating of young thinkers, which is blinder than any other self-conceit (*Eigendiünkel*) and more incurable than ignorance'.⁸ A student who exceeds his boundaries will know the language of science, but lack the ability to use it correctly or critically. This is a recurring theme of self-conceit: the conceited agent wields the language of a discipline for his own purposes, rather than for the sake of discovery or critical analysis. Of course, the precocious child is more careless than selfish; presumably he uses such language to seem sophisticated, not necessarily to win an argument. Even so, a troublesome effect begins to emerge from this precociousness. Specifically, Kant observes, a pupil becomes difficult to educate or correct. Ignorance can be corrected via the application of standards and analysis, but the precocious student has already appropriated a mock version of these for his own purposes.

A more self-interested form of overstepping occurs within philosophical and scientific discourse. In his discussion of the ideal of reason in the *Critique of Pure Reason*, Kant notes that 'physico-theologians' look down on the transcendental proof with 'the self-conceit (*Eigendiünkel*) of clearsighted students of nature', but in their endeavours they 'leave this territory of nature and experience and pass over into the realm of mere possibilities, where on the wings of ideas they hope to approach that which has eluded all their empirical enquiries' (Kant 1998: 583; *KrV*, A629–30/B657–68). Here, the 'overstepping' (*Übersteigen*) associated with self-conceit manifests itself in a tendency to supplant what is missing after empirical enquiry has been exhausted with imagined ideas and possibility. This can take the form of reason simply getting ahead of

itself: Kant thinks that we sometimes have a tendency to ascribe an efficient cause too hastily.⁹ But, more often and more seriously, this self-conceit of overstepping takes the form of thinking that one can intuit something that could not possibly be the object of experience, but that one wishes, nevertheless, to be true. The ‘overstepping’ associated with self-conceit in such a case is not just an instance of careless excitement, as it was for the precocious student. Many of Kant’s references to the failure of self-conceit in theoretical reason occur in the context of discussions about scientific and philosophical debate.¹⁰ When, in these contexts, a person supplants what is missing with imagination and possibility, she is doing so in an attempt to prove a point or win an argument.¹¹

In the practical sphere, the overstepping associated with self-conceit manifests itself as moral fanaticism, or *Schwärmerei*.¹² Kant explicitly links self-conceit and fanaticism in the second *Critique* (*KpV*, 5: 84), and characterizes fanaticism as a failure that involves overstepping the bounds of reason:

If fanaticism (*Schwärmerei*) in the most general sense is an overstepping (*Überschreitung*) of the bounds of human reason undertaken on principles, then *moral fanaticism* (*moralische Schwärmerei*) is such an overstepping of the bounds that pure practical reason sets to humanity, thereby forbidding us to place the subjective determining ground of dutiful actions – that is, their moral motive – anywhere else than in the law. (Kant 1996a: 209; *KpV*, 5: 85–6)

Reason and common sense confirm that the ground of obligation can be found only in the moral law, which, in turn, cannot be grounded in experience. The mistake of fanaticism is to overstep this restriction and imagine that obligation could be grounded in experience. The presumption that duty is based in experience is, on Kant’s view, far more fantastical than the acknowledgement of its grounding in pure practical reason. As the *Groundwork* puts it, ‘one cannot better serve the wishes of those who ridicule all morality as the mere phantasm of a human imagination overreaching itself through self-conceit, than by conceding to them that the concepts of duty had to be drawn solely from experience’ (Kant 2011a: 43; *G*, 4: 407).

The moral fanatic need not commit this error of empiricism deliberately or with malice aforethought. A major culprit of *Schwärmerei* is the tendency to imitate moral action, rather than subject oneself to the moral

law (*KpV*, 5: 85). Note here an analogy with the precocious student cited above: in imitating the language of science, this student became ‘incorrigible’ and unable to use the standards of theoretical reasoning critically. Moral imitation has the same effect of untethering the agent from the representation of duty and instead leading to a ‘frivolous, high-flown, fantastic cast of mind’ (Kant 1996a: 208; *KpV*, 5: 85). This cast of mind gives rise to the most salient feature of moral fanaticism – the tendency to replace respect for the moral law with self-congratulatory, pathological sentiment. This result stands to reason: respect for the moral law can only come about when one’s self-love and self-conceit are constrained or thwarted by a recognition of the moral law, and when one experiences the humiliation that results from this encounter (*KpV*, 5: 74). Lacking any such encounter, the fanatic replaces respect with pathological sentiment and is able to imagine that his inclinations reflect unfailingly the demands of morality. Kant describes the conceited moral agent as one who allows his thoughts to ‘rove among fancied moral perfections’ (Kant 1996a: 209; *KpV*, 5: 86). Such fancied perfections are decidedly not strivings after a perfectly good will; rather, their object is the ‘unreachable perfection’ that the ‘mere heroes of romance’ evince in the Doctrine of Method of the second *Critique* (Kant 1996a: 264; *KpV*, 5: 155). Theirs is the imagined perfection of having one’s inclinations align effortlessly with morality.

In opting for pathological fanaticism over the representation of duty, the conceited agent opts for a form of heteronomy that includes the delusion that his sentiments and inclinations are adequate and accurate guides to moral action. The conceited agent demonstrates an interest in appearing – to himself and to others – as though he has a virtuous character. But he has no interest in participating in the moral struggle and sacrifice associated with autonomous willing. Thus, rather than acknowledge that his inclinations can and should be subordinated to a moral law whose authority reason recognizes, the conceited agent constructs an illusion about the native perfection of his empirical will in the face of temptation or moral danger.¹³ The conceited agent fails to recognize himself as the subject of moral striving, and that the ‘proper moral condition, in which he can always be, is *virtue*, that is, moral disposition in conflict, and not holiness in the supposed possession of a complete purity of dispositions of the will’ (Kant 1996a: 208; *KpV*, 5: 84).

In its claims to effortless perfection, moral self-conceit is thus the opposite of timorousness (*Muthlosigkeit*) and despondency (*Kleinmüthigkeit*), notions that Kant explicitly contrasts with self-conceit in his lectures on

moral philosophy.¹⁴ Despondency occurs when, upon realizing his moral shortcomings, an agent decides that none of his actions could ever conform with the moral law, a conclusion that leads to a type of resigned indolence. As the stated antithesis of self-conceit, despondency is an informative concept. Notably, despondency is not necessarily a judgement about the relative moral *superiority* of others. It is, rather, a judgement about one's own moral *ineptitude*. Despondency is the mistaken judgement that one can never live up to the moral law, while self-conceit involves the opposite, yet also mistaken, judgement that one cannot *help but* live up to the demands of morality. Both the conceited agent and the despondent agent misinterpret their own moral shortcomings. The conceited agent recasts his failures, via fanaticism and rationalization, as moral successes. The despondent agent, on the other hand, takes his shortcomings to be evidence that he will never be adequate to the moral law. Notably, however, both errors stem from the failure to recognize that virtue consists in a continuous struggle. In the case of the conceited agent, this mistaken assumption manifests itself in the agent's enthusiastic delusion that she is effortlessly virtuous. The despondent agent interprets moral failure not as a stumbling block that requires overcoming, but as evidence that his is not a will endowed with natural perfection. The despondent agent is thus analogous to the child who gives up on a project or assignment when, having always been told that he is 'naturally gifted', he encounters his first challenge or criticism.

Solipsism and imposing judgements upon others

One problematic consequence of the overstepping associated with self-conceit is that the conceited agent takes the judgements of others to be unnecessary.¹⁵ In his lectures on logic, Kant identifies this failure with 'logical egoism' and argues that it consists in 'the presumed but often false self-sufficiency of our understanding ... where one believes he knows enough by himself, and believes he is infallibly correct and incorrigible in all his judgements' (Kant 1992: 119; *VL-Blomberg*, 24: 151).

We may wonder about Kant's warning against taking one's own understanding to be self-sufficient. It would be a curious result, indeed, if Kant were to assert that theoretical or practical reason depended on others for the cognition of truth. But this is not Kant's claim. Rather, Kant is aware of the fact that we often make errors in our reasoning: theoretical reason tends to fall confusedly into antinomies, and pure practical reason is led astray by inclination and subsequent rationalization. Kant's suggested solution to this danger is to recommend a principle of publicity. Accordingly, Kant makes a distinction among ways of 'taking something

to be true' (*Fürwahrhalten*). Among other things, Kant introduces a distinction between mere persuasion (*Überredung*) and conviction (*Überzeugung*). Persuasion, Kant explains, is an instance of taking-to-be-true that 'has its ground only in the particular constitution of the subject' (Kant, 1998: 685; *KrV*, A820/B848). Problematically, however, a person whose assent is mere persuasion also takes her judgement to have objective grounds.¹⁶ Kant explains that there is a straightforward test (or 'touchstone') that can be employed to distinguish between instances of mere persuasion and conviction, namely, whether a person can communicate an instance of taking-to-be-true and find it 'to be valid for the reason of every human being to take it to be true' (*ibid.*). Kant's point is thus not that cognition of the truth depends on intersubjective agreement; it is, rather, a point about self-knowledge and verification. If a person's reasons for taking something to be true are incommunicable, this is a sign that hers is mere persuasion, not conviction.¹⁷

The conceited agent thus takes the judgements of others to be unnecessary in just this sense: she sees no possible need for intersubjective verification as a touchstone against error. As far as the conceited agent is concerned, she *cannot* make mistakes in her reasoning. A contributing factor to this error is the conceited agent's tendency to misinterpret evidence, and Kant associates self-conceit with derangement (*Wahnsinn*) or dementia (*Verrückung*).¹⁸ Kant describes dementia as the state that occurs when imagination fools the senses into seeing something that could never be the object of experience (*VAnth*, 25: 1008). The demented person has the *capacity* to experience things correctly, but *interprets* experience, especially his experience of other human beings, 'through an absurd delusion as referring to himself' (Kant 2008b: 74; *VK*, 2: 268).

In the context of practical reason, this tendency to interpret evidence to one's own liking manifests itself as a failure of self-knowledge and self-evaluation. One instance of this tendency is the conceited agent's propensity to take moral credit where it ought not be taken, particularly in cases of moral luck. Kant wonders:

if we subtract from the causes of a presumably well-led course of life everything which is called the merit of fortune ... whether before the all-seeing eye of a world-judge one human being has any superiority over another regarding his moral worth? And, on the basis of this superficial self-knowledge, might it not perhaps be absurd self-conceit (*ungereimter Eigendünkel*) to pronounce any judgement at all to one's own advantage concerning one's own

moral worth or that of others (or the fates they deserve)? (Kant 1996b: 223; (*EaD*), 8: 330)

Similarly, in his discussion of malice in the Doctrine of Virtue, Kant observes that, when it comes to malevolent individuals, it is in part ‘their self-conceit (*Eigendiünkel*) in their good conduct (strictly speaking, only their good fortune in having so far escaped temptations to public vice) ... which an egotist accounts to his merit’ (Kant 1996c: 577; *MS*, 6: 460). The conceited agent thus counts mere accidents of fortune, such as a kind temperament or having escaped temptation, towards a positive assessment of his character or virtue.

But the failure of self-conceit does not end in the conceited agent’s being indifferent to the judgements of others. Rather, he often ultimately *imposes* his judgements upon others. The lectures on logic thus sometimes describe self-conceit as a *subspecies* of logical egoism in which the agent takes it upon himself to make his unfounded persuasion the basis of a universal judgement.¹⁹ This characterization of self-conceit naturally draws to mind the analyses of self-conceit discussed above in section 2, according to which self-conceit is a kind of imperiousness or arrogance. Again, there is no question that these attitudes are associated with Kantian self-conceit, but Kant’s remarks in the logic lectures demonstrate a crucial point about these attitudes. Specifically, they are a straightforward *result* of the attitudes about reasoning described above, and not themselves the ultimate source of self-conceit. The conceited agent’s first concern is with getting her way, whether in terms of theoretical judgement or practical conclusion. But rather than opt out of the standards of, and constraints upon, reasoning that otherwise prevent her from always getting her way, she appropriates and manipulates these for her own purposes – by overstepping the bounds of reason with imagination, misinterpreting evidence, or constructing an illusion of native perfection. Ultimately, in distorting these principles of reasoning for her own aims, she also makes an implicit judgement about other agents – that their judgements as ‘touchstones’ are superfluous. Further, in order to maintain this illusion of objective validity, she is compelled to impose her judgements upon others. The fact that attitudes of arrogance and judgements of worthlessness are the *result* of these failures of reasoning is precisely what makes the failure of self-conceit so intractable and difficult. It is not enough to remind the conceited agent of the standards of rationality or the equality of agents. The conceited agent takes herself to be recognizing these facts. However, she misappropriates a reliable test for error – a principle of publicity – for her own aims when she imposes

her judgements upon others in order to maintain the delusion that she is abiding by the very constraints to which she makes an exception.

In the practical sphere, self-conceit is less an *inclination*, and more precisely a mode of thinking about one's inclinations and moral choices.²⁰ Notably, Kant refers to self-conceit as a *principle* of the will in the second *Critique*, whereas self-love is described as a determining *ground* of the will. Self-conceit, on this account, does not motivate directly; rather, it fortifies and legitimizes the choices suggested by the faculty of desire by recasting them as principled actions. Through rationalization and the construction of a delusion about her native perfection, the conceited agent rationalizes her heteronomy into virtue. As Kant's early lectures put it, self-conceit involves 'tinkering' (*künsteln*) with the moral law (Kant 1997: 216; *VE-Collins*, 27: 465). This feature of self-conceit is informative in two respects. First, the object of the conceited agent's concern is *herself*, specifically, the estimation of her own moral worth. The conceited agent does not seem to be preoccupied, in the first instance, with the worthlessness of others. Second, the conceited agent obviously takes an interest in appearing as though she is acting from duty, and in attaining the moral approval and respect of others (this will become especially apparent in the next section). Of course, the conceited agent's concern for the moral law does not extend to having *respect* for the law. Nevertheless, she recognizes some value in the appearance of moral worth, even if she has little interest in doing the difficult work required in achieving such worth.

4. Self-Love and Self-Conceit

Having examined the habits and tendencies of the conceited agent, we are in a better position to consider Kant's explicit analyses of moral self-conceit, which often appear in the context of a comparison with self-love. This section considers several discussions in which Kant glosses the distinction between self-love and self-conceit, beginning with the lectures on moral philosophy ascribed to Johann Friedrich Kaehler and Georg Ludwig Collins, and concluding with the discussion of the terms in the Doctrine of Virtue. I reconsider Kant's discussion of self-love and self-conceit in the *Critique of Practical Reason* in the concluding section of the paper.

We begin with the discussion of self-love and self-conceit as it appears in lectures on moral philosophy from the mid to late 1770s.²¹ As Kant puts the distinction there:

[T]he love that takes pleasure in oneself (*Die Liebe des Wohlgefallens gegen sich selbst*), a self-love, is an inclination to be well-content with

oneself in judging of one's perfection. *Philautia*, or moral self-love, is to be contrasted with arrogance, or moral self-conceit (*Eigendiinkel*). The difference between them is that the former is only an inclination to be content with one's perfections, whereas the latter makes an unwarranted pretension to merit. It lays claim to more moral perfections than are due to it; but self-love makes no demands, it is always merely content with itself and devoid of reproach. (Kant 1997: 135; *VE-Collins*, 27: 357)

This passage is puzzling on its face. Curiously, self-love appears here to overlap in important respects with Kant's gloss of self-conceit in other texts. For example, Kant describes self-love as a kind of *liking of oneself* (*Wohlgefallens gegen sich selbst*), whereas the language of well-liking, or *Wohlgefallen*, is typically reserved for self-conceit elsewhere.²² Further, in this passage, self-love is described as an inclination to be 'well-content with oneself in judging one's perfections', and this sounds like the account of fanaticism and self-deception offered provisionally in the preceding section of this paper. But if this is how Kant understands self-love, it becomes even more difficult to discern a distinction between self-love and self-conceit.

However, the passage from the Kaehler/Collins lecture is best interpreted alongside the corresponding sections of Alexander Baumgarten's textbook, *Ethica Philosophica*, which Kant used for his courses on moral philosophy. In his own discussion of self-love, Baumgarten makes a distinction between well-ordered self-love (*philautia ordinata*, glossed in a footnote as 'wohlgeordnete' *philautia*) and disorderly self-love (*philautia inordinata*, glossed as 'unordentliche' *philautia*) (Baumgarten 1969: §194). Well-ordered self-love, according to Baumgarten, is a self-love that responds with appropriate approval to an accurate assessment of one's moral perfections. Disorderly self-love, on the other hand, is 'blind' and imagines false perfections. In light of Baumgarten's distinction, Kant seems simply to be making a similar point. Accordingly, self-love in the Kaehler/Collins passage is analogous to Baumgarten's orderly self-love, and is the inclination to be content with the 'perfections' that one legitimately recognizes in oneself. Self-conceit, on the other hand, is analogous to Baumgarten's disorderly self-love, and is the tendency to make unwarranted 'pretension(s) to merit'. As we will see in what follows, though Kant's notion of self-love shifts in later works, this conception of self-conceit remains consistent.²³

Another explicit distinction between self-love and self-conceit does not emerge until the second *Critique*. But, of course, Kant discusses self-love in

texts that precede that work, and in these the concept famously takes on a sense different from the sense presented in the Kaehler/Collins lectures. In the *Groundwork*, self-love operates as an incentive, i.e. as ‘the subjective ground of desiring’ (Kant 2011A: 83; G, 4: 427). Self-love has this function, for example, in the various maxims that are tested according to the formula of universal law. In the *Groundwork*, then, self-love is an incentive generated by the faculty of desire, not an assessment of one’s own virtue. Similarly, in the roughly contemporaneous lectures on moral philosophy transcribed by Christoph Mrongrovius, Kant notes that (the ultimately rejected) ‘inner empirical grounds’ for morality can be based in either self-love or moral feeling. The former, Kant reportedly explains, is ‘the principle of the feeling of what is good for us, or the physical feeling, [and] is the principle of happiness’ (VE-Mrongrovius II, 29: 622–3). Again, Kant associates self-love with the faculty of desire, specifically with our natural desire for happiness.

Once self-love takes on this sense, the discussion of self-love and self-conceit no longer tracks only one variable, i.e. the extent to which one correctly assesses one’s moral perfections. Instead, the distinction between self-love and self-conceit begins to track two distinct questions or variables: questions regarding self-love concern an agent’s desires and inclinations, and the extent to which these are decisive in action and moral choice. Questions of self-conceit concern the agent’s *assessment* of her own maxims and choices. Notably, in the Vigilantius lectures on moral philosophy (1793/94), Kant points to the same passage in Baumgarten’s text, and amends Baumgarten’s account of *philautia*:

Our author has not adequately distinguished *philautia* from two other competing concepts. There are people who see themselves as objects of general love in relation to others, and so also behave to others in such a way as to demand general love and respect from them. *Philautia* therefore rests on having a good opinion of ourselves; only with this difference, that in doing so we either regard ourselves as a subject worthy of love (*ein lebenswürdiges Subject*), and this is *philautia*, or as one worthy of respect (*ein achtungswürdiges*), and this is arrogance. (Kant 1997: 363; VE-Vigilantius, 27: 620)

Kant begins again with Baumgarten’s analysis, namely that *philautia* involves ‘having a good opinion of ourselves’. But now Kant observes that this description is incomplete: *philautia* can manifest itself as the ‘demand’ for love, i.e. as the thought that one is worthy of more than

one's fair share of love from others. But *philautia* can also manifest itself as a 'demand' for respect, i.e. as the thought that one is worthy of more than one's fair share of respect from others. The former is self-love proper; the latter is arrogance (*Arroganz*). Though thinking of oneself as worthy of love and respect is certainly acceptable in some cases, Kant's use of the pejorative word 'arrogance', read alongside the suggestion that these attitudes involve 'demands', suggests that this passage is concerned with blameworthy instances of these attitudes.

It should also be noted that Kant does not use the term self-conceit (*Eigendünkel*) in this passage; however the parallels between this passage and other passages in which self-love and self-conceit are contrasted suggest a common referent. For example, in the longer discussion that follows, Kant reportedly describes self-love as 'the love of well-wishing towards oneself' (*die Liebe des Wohlwollens gegen sich selbst*) (Kant 1997: 363; *VE-Vigilantius*, 27: 620) and arrogance as 'the love of well-liking toward oneself' (*die Liebe des Wohlgefallens gegen sich selbst*) (Kant 1997: 364; *VE-Vigilantius*, 27: 621). Recall that this is precisely the distinction between self-love and self-conceit that Kant offers in the *Critique of Practical Reason* (*KpV*, 5: 73).

In the *Vigilantius* lecture, Kant thus classifies self-love (*die Liebe des Wohlwollens gegen sich selbst*) and arrogance (*die Liebe des Wohlgefallens gegen sich selbst*) as subspecies of a broader, umbrella-notion of self-love (*philautia*).²⁴ This is a regular pattern in Kant's later work. In a note on the use of the term 'self-love' (*Selbstliebe*) in *Religion Within the Boundaries of Mere Reason*, Kant remarks that 'love' is a word that can be taken to have two entirely different meanings (*RGV*, 6: 45n.). Both 'love' and 'self-love' can be understood as *either* the love of well-wishing (*Wohlwollen*) or as the love of well-liking (*Wohlgefallen*).²⁵ The love of well-wishing takes as its object either our own happiness and ends, or the happiness and ends of others.²⁶ Because we all desire our own happiness, it is 'natural', according to Kant, to incorporate the self-love of well-wishing into one's maxims, though it ought never become the unconditional principle of the will. When it comes to the self-love of well-liking, Kant explains that this term can be understood in two further senses. First, well-liking of oneself can be an instance of being satisfied that one has successfully carried out a hypothetical imperative: the merchant whose speculations turn out well has an attitude of well-liking towards himself in response to the maxims that he wisely chose to act upon. The second form of well-liking, at issue in this discussion, is moral well-liking, or contentment (cf. *KpV*, 5: 117). Again, neither the self-love

of well-wishing, nor the self-love of moral well-liking is, in itself, problematic. To see how these become problematic, we return to the discussion in the *Vigilantius* lecture.

The *Vigilantius* lecture describes blameworthy instances of well-wishing and well-liking of oneself as forms of solipsism. Such solipsism is most straightforwardly evinced in the case of blameworthy self-love (*VE-Vigilantius*, 27: 620). In the *Vigilantius* lecture, Kant defines ‘love’ as the inclination or will to promote the ends of others. *Self-love*, on this account, would be the inclination or will to promote one’s own ends. *Prima facie*, there is nothing necessarily blameworthy in the inclination or will to promote one’s own ends. However, self-love becomes blameworthy when it becomes solipsistic, or when it ‘excludes others from our love or inclination toward them’.²⁷

Arrogance (or self-conceit) is the blameworthy variety of well-liking of oneself, and this, too, is described in terms of a kind of solipsism in the *Vigilantius* lecture.²⁸ This claim is less straightforward on its face than the claim that blameworthy self-love is solipsistic. If well-liking concerns the assessment of one’s moral worth (*RGV*, 6: 45n.), then it is not entirely obvious how this assessment should require others. Moral agents ought to compare themselves to the moral law, not others, in order to determine the worth of their actions.²⁹ But as the above discussion of self-conceit in the logic lectures has shown, Kant’s claim is not that we require a comparison with others in order to settle the question of moral worth. Rather, other agents serve as moral touchstones who can help us correctly interpret the moral worth of our maxims and choices. The agent whose well-liking is solipsistic sees no need for others as moral touchstones; she would see no need, for example, to engage in a moral friendship with another agent who might point out to her various errors in her reasoning (*VE-Vigilantius*, 27: 426–8).

Unsurprisingly, Kant remarks that the agent who evinces this sort of solipsistic well-liking ‘puts himself in danger of being incapable of examining or amending his faults’ (Kant 1997: 364; *VE-Vigilantius*, 27: 621). This follows straightforwardly from what we have already seen: an agent whose self-assessment turns inward and excludes moral touchstones will run the risk of moral delusion or fanaticism; he will suffer from a kind of persuasion, or *Überredung*, with respect to the moral worth of his actions. This is worrisome enough. But in the *Vigilantius* lecture, there is a hint about how this failure might develop into a fanaticism that is ultimately imposed upon others. Kant remarks that well-liking

towards oneself, in comparison with others, is ‘differentiated and transformed into self-estimation of oneself’ (*Selbstschätzung seiner selbst*) (ibid.). The point is best understood as a claim about the respect that we merit from others as a function of our status as moral agents: Kant remarks that, if this self-estimation ‘rests on *a priori* close examination of oneself’, it is ‘self-justifying in its own right’ (ibid.). However, if ‘it rests, without examination, on a judgement of oneself, whereby the agent makes himself an object of the respect that we require from others, and enhances his worth, without justification, over that of other people, then it is arrogance, and a fault’ (ibid.).

As far as our relations with others are concerned, self-conceit manifests itself as a demand for respect from others that results from a delusion regarding our own virtue. At times Kant suggests that it is, strictly speaking, the moral law and not individual agents that we ought to respect (e.g. *G*, 4: 401n.). But often enough, Kant gives us examples of individuals whose moral conduct generates in us a respect for them, and for the moral law (e.g. *KpV*, 5: 77 and 156). Thus, to ‘make oneself an object of ... respect ... without justification’ is to expect others to accord one more respect than one actually deserves for acting from duty. Note further that the conceited agent does not simply want to be seen as avoiding moral error, or abiding by perfect duty. Presumably, she wants to be respected and praised for some positive virtue, when in fact she has only acted from inclination and rationalized her actions in a delusion of moral perfection and fanaticism. Importantly, none of these failures is, in the first instance, a judgement about *others*, or even a judgement about one’s worth relative to others. As far as the conceited agent is concerned, she is being judged by precisely the same moral standard as everyone else; she simply has the great fortune of possessing an effortless and native virtue that warrants respect according to those standards. The conceited agent’s failure is thus at bottom a failure of self-knowledge and self-evaluation, which she then imposes upon others in order to preserve the illusion of virtue. In a discussion of arrogance and self-respect in Kant’s moral philosophy, Robin Dillon incisively identifies self-conceit with this set of failures and cites Kant’s discussion of ‘moral arrogance’ in the Doctrine of Virtue as a straightforward statement of the problem. As Kant puts it, ‘moral arrogance’ (*Tugendstolz*) is a ‘conviction of the greatness of one’s moral worth, but only from a failure to compare it with the law’ (Kant 1996c: 558; *MS*, 6: 435). Dillon dubs this the failure of ‘primary arrogance’, which she defines as ‘an unwarranted claim to much more moral merit than one has actually earned’.³⁰

Kant's analysis of self-love and self-conceit in the Doctrine of Virtue suggests a similar account:

Moderation in one's demands generally, that is, willing restriction of one's self-love in view of the self-love of others (*freiwillige Einschränkung der Selbstliebe eines Menschen durch die Selbstliebe Anderer*), is called modesty. Lack of such moderation (lack of modesty) as regards one's worthiness to be loved by others (*in Ansehung der Würdigkeit von Anderen geliebt zu werden*) is called egotism (*Eigenliebe [philautia]*). But lack of modesty in one's claims to be respected by others (*von Anderen geachtet zu werden*) is self-conceit (*arrogantia*). The respect that I have for others or that another can require from me (*observantia aliis praestanda*) is therefore recognition of a dignity (*dignitas*) in other human beings, that is, of a worth that has no price, no equivalent for which the object evaluated (*aestimii*) could be exchanged. – Judging someone to be worthless is contempt. (Kant 1996c: 579; *MS*, 6: 462)

First, a note on the translation: though the English word 'egotism' is very similar to 'egoism' (*egoismus*), the German word that Kant uses in this passage is *Eigenliebe*, and Kant adds the Latin *philautia* as further clarification. These two words appear together consistently in Kant's discussions of self-love. Most notably, Kant uses both terms (*Eigenliebe* and *philautia*) as terms for self-love in the crucial discussion of self-love and self-conceit in the second *Critique*, discussed at the outset of this paper. In short, we should not be misled by the use of 'egotism' in the English translation. In this passage, Kant is clearly dealing with the distinction between self-love (*Eigenliebe*, *philautia*) and self-conceit (*Eigendünkel*, *arrogantia*).

The passage shares some obvious similarities with the Vigilantius discussion of self-love and self-conceit. First, Kant's classification scheme remains consistent: self-love and self-conceit are here subspecies of a broader notion of self-love. This classification is also consistent with Kant's remarks on these terms in the *Religion*. (However, as we have already seen, Kant uses the term *Selbstsucht* (self-regard) as the umbrella-term in the second *Critique*.) Second, while Kant does not use the terms *Wohlwollen* and *Wohlgefallen* here, the wish to be 'loved by others' and the wish to be 'respected by others' parallel the notions of well-wishing (*Wohlwollen*) for oneself and well-liking (*Wohlgefallen*) of oneself, respectively.

However, Kant describes both self-love and self-conceit in terms of an agent's relations and attitudes *towards others*, specifically in terms of his judgement that he is worthy of the love of others, and his judgement that he has a claim to their respect. Indeed, even in the *Vigilantius* passage above, Kant describes the unreasonable agent's solipsism (*Solipsismus*) in terms of his 'demands' and his judgements that he is worthy of the love and respect of others. What shall we make of these assertions? Are they evidence that (unreasonable) self-love and self-conceit are fundamentally failures of arrogance? Not necessarily. Let us take the claims of unreasonable self-love first. Every agent desires his own happiness, and, to the extent that his happiness depends on others adopting his ends, he wills (or, more accurately, wishes) that others will adopt his ends. This wish that others adopt one's ends is acceptable, so long as it is consistent with the demands of the moral law. The concern for one's happiness becomes problematic only when this willing becomes solipsistic (*VE-Vigilantius*, 27: 620), or when an agent replaces the 'objective determining ground of the will in general' (the moral law) with 'subjective determining grounds of choice' (Kant 1996a: 199; *KpV*, 5: 73). When an agent does this, his hope that others will adopt his ends turns into an *expectation* that they will adopt his ends, or at least a judgement that others should adopt his ends. To replace objective grounds with subjective grounds is thus to view one's own happiness as supremely governing for oneself *and others*, insofar as their decisions to adopt one's ends relate to one's happiness.³¹ It is thus no surprise that Kant equates the solipsism of unreasonable self-love with the agent's judgement that he has a legitimate claim to others' love of well-wishing.

But Kant also mentions respect and *dignity* in this passage, suggesting not only that the conceited agent demands unwarranted respect from others, but also that he fails to recognize a dignity in others, thus evincing contempt for them. On their face, these claims suggest an interpretation of self-conceit that involves arrogance and judgements of worthlessness at its very core. However, this is to imbue Kant's terminology with an anachronistic interpretation. Crucially, Kantian claims about dignity are relational claims: they are assertions about the shared status of rational human agents, namely, that rational agents are set above and apart from other creatures in virtue of their freedom.³² Accordingly, to demand recognition of dignity while denying others such recognition amounts to an assertion that one is set apart from others because of a capacity for freedom that others lack. A failure to recognize dignity in this respect is a failure to see others as free and rational agents with whom one must share justificatory space by, for example, universalizing maxims or seeking consent.³³ Note, too, that the conceited agent's failure to ascribe dignity

to others corresponds well with the account of theoretical self-conceit sketched earlier. The conceited scientist takes her own imagined evidence and investigation to be objectively valid. Kant observes that the conceited scientist misjudges his own powers and is a kind of ‘Cyclops’ who ‘needs another eye’. Kant thus concludes that the ability to see one’s subject matter from another point of view grounds the ‘humanity of the sciences’ (*Refl.* 15: 394). Similarly, the conceited moral agent makes a judgement about *her own* practical reason, namely, that she need not universalize her maxims or seek consent from other agents who share her status as a free, rational agent. Instead, she simply pursues her own ends and expects to be respected or praised for the effortless moral worth of her choices.

The core failure of self-conceit is not a judgement of the worthlessness of others, but rather a solipsism and self-congratulation that excludes others, both as moral touchstones and as agents to whom one ought to be able to justify oneself. Self-conceit is thus fundamentally a failure of one’s duties to oneself. Indeed, in a note in his drafts for the *Metaphysics of Morals*, Kant remarks that there are two duties – ‘1. to make the ends of others my own [and] 2. To *lower my claims to self-conceit* through the absolute right that every other has as a human being (that of equality)’.³⁴ In the *Metaphysics of Morals* itself, Kant ultimately settles on two ends that are at the same time duties – the happiness of others, understood as adopting the ends of others as one’s own, and one’s own perfection, understood as perfecting one’s virtue (*MS*, 6: 385). The association between ‘lowering one’s claims to self-conceit’ (in the draft) and perfecting one’s virtue (in the published text) is informative. The agent who makes an effort to avoid self-conceit recognizes the moral law as supremely binding and makes an effort to avoid the tendencies toward rationalization and self-deception that would otherwise obscure instances of heteronomy or moral failure. Kant’s claim that we have a duty to ourselves to ‘lower [our] claims to self-conceit’ thus reinforces his assertion that duties to oneself are primary, and a condition for the performance of all other duties (*MS*, 6: 417–18). An agent cannot act from duty unless she sees herself as a moral agent, bound by the moral law and striving to live up to its demands. Any other attitude admits, at best, of merely accidental coherence with the moral law. ‘Lowering one’s claims to self-conceit’ amounts to the recognition of the moral law that an agent must give to herself if she is to will and act upon moral maxims at all.³⁵

5. Conclusion: Striking Down Self-Conceit

We now return to Kant’s discussion of self-conceit in the *Critique of Practical Reason*. Recall that there Kant describes unreasonable self-love

as the tendency to turn ‘subjective determining grounds of choice into the objective determining ground of the will in general’. Self-conceit is the tendency to make self-love ‘itself lawgiving and the unconditional practical principle’ (Kant 1996a: 200; *KpV*, 5: 74). The moral law ‘infringes’ upon self-love, but ‘strikes down’ self-conceit (Kant 1996a: 199; *KpV*, 5: 73). Let us begin with this last point: why does the moral law ‘strike down’ self-conceit, but only ‘infringe upon’ self-love?

The answer is straightforward, in light of the preceding discussion. Self-love, or well-wishing for oneself, can be reasonable, insofar as it is constrained by the moral law. However, Kant argues that the moral law ‘strikes down self-conceit altogether, since all claims to esteem (*Selbstschätzung*) for oneself that precede accord with the moral law are null and quite unwarranted because certainty of a disposition in accord with this law is the first condition of any worth of a person’ (ibid.). There are, to be sure, legitimate forms of *esteem* for oneself, but self-conceit is always an illegitimate form of esteem. It is the delusion that one is perfectly virtuous, fuelled by a failure of self-knowledge and self-examination. As such, it is simultaneously a *de facto* denial of the unconditionally binding force of the moral law and a lie that one tells oneself about the moral worth of one’s actions. There is thus no ‘reasonable’ type of self-conceit. *All* of its esteem precedes ‘certainty of a disposition in accord with [the] law’, because its esteem is a reverence for fictional virtue.

How, then, shall we understand Kant’s claim that self-conceit is the tendency to make self-love ‘itself lawgiving and the unconditional practical principle’? First, we should resist the conclusion that self-conceit necessarily involves *imposing* one’s unreasonable self-love upon others as a matter of principle. As argued above, any unreasonable form of self-love, i.e. self-love that becomes the ‘objective determining ground of the will’, imposes itself upon others insofar as the agent ‘demands’ that others aid him in his quest for happiness. More to the point, there is nothing about self-conceit that is essentially or necessarily other-directed. It is entirely possible, for example, that an agent living in isolation could make the errors associated with self-conceit. This conceited agent might convince herself that when she acts on an inclination to take long afternoon naps, she is actually fulfilling a duty to herself to look after her health. Now, of course, most agents do not live in isolation, so when they fall into self-conceit, their rationalizing and self-congratulatory attitudes lead them also to impose these illusions of virtue upon others. However, the fundamental failure of self-conceit is a failure in the way an agent assesses herself morally. Through self-deception and delusion the fanatic

disguises heteronomy as autonomy. She disregards the real unconditional practical principle because she has constructed a comfortable and convincing substitute based on her own inclination. Thus, self-conceit does not make self-love the unconditional practical principle in the sense that it *imposes* an agent's self-love upon others. Rather, self-conceit makes self-love the unconditional practical principle in the sense that it reconstructs an agent's moral self-conception, including her capacity for moral self-evaluation, around self-love and heteronomy.

What, then, of the second set of questions raised at the outset of this paper? First, can the failure of unreasonable self-love occur absent self-conceit? Self-conceit does seem a probable by-product of unreasonable self-love on the Kantian account. Common sense recognizes the moral law and finds it compelling and worthy of reverence, even if this means, in a perverted sense, that an agent might only pretend to follow it. Still, self-love need not rely upon or generate self-conceit. Numerous characters and concepts in Kant's moral thought demonstrate that an honest assessment of one's wrongdoing is possible, even if it is difficult or painful. A case in point is Kant's discussion of conscience and self-knowledge in the Doctrine of Virtue.³⁶ Conscience, Kant argues, is one's 'consciousness of an *internal court* in the human being' (Kant, 1996c: 560; *MS*, 6: 438). It is not a judgement that one 'voluntarily ... makes', but rather an awareness that is 'incorporated into [one's] being' (*ibid*). Conscience is thus an internal authority that compels an agent to submit his actions to the scrutiny of self-evaluation and judgement. Especially important to the question at hand is Kant's claim that conscience is *ever-present*: an agent 'can at most, in extreme depravity, bring himself to *heed* it no longer, but he still cannot help *hearing* it' (*ibid.*). On Kant's view, then, even the thoroughly conceited agent will 'hear' the infallible voice of conscience. This observation may, in turn, assuage any worry that the thoroughly deluded agent can no longer be held responsible for her wrongdoings.

Accurately knowing oneself 'in terms of [one's] moral perfection in relation to [one's] duty' (Kant 1996c: 562; *MS*, 6: 441) is thus to make an accurate verdict of oneself in the tribunal that conscience initiates. Such cognition is painful, but possible: we have a duty to achieve it. Further, this moral cognition guards against precisely those two errors of fanaticism discussed above in section 3. Unsurprisingly, moral cognition 'counteract[s] that egotistical self-esteem which takes mere wishes ... for proof of a good heart'. But moral cognition also dispels despondency, or a 'fanatical contempt for oneself' since it makes the agent aware of a 'noble predisposition to the good' in himself and other human beings,

even if he is aware of having fallen short of the moral law.³⁷ Again, Kant's claim is evidence that self-conceit and despondency are close bedfellows. Both deny the unconditionally binding force of the moral law by constructing an illusion of native perfection. This illusion is, in both cases, dispelled by accurate self-knowledge and the recognition of virtuous struggle.

Self-love is possible absent self-conceit, but what of the converse? Can self-conceit exist absent unreasonable self-love? On its face, this would seem to be impossible. The decision to subordinate the moral law to self-love seems at the very least to provide the *occasion* for self-conceit. But, in fact, unreasonable self-love does not provide the *only* occasion for self-conceit. In the *Vigilantius* lecture, Kant observes that self-conceit (arrogance, 'unreasonable well-liking toward oneself') can occur:

without any selfishness (*Eigennutz*), insofar as we harbour a high degree of well-wishing towards others. It is true, however, that owing to it there commonly arises the source of self-complacency in its defective form, in that by unselfishly promoting the welfare of others we engender a self-satisfaction in ourselves, and respect ourselves self-lovingly, without estimation of our true moral worth. If this self-satisfaction is to be true self-esteem, we have to distinguish between the performance of obligatory and meritorious acts of duty. Obligatory duties toward others are indeed a mere settlement with them, without our being able to demand anything from them on that account. (Kant 1997: 364; *VE-Vigilantius*, 27: 621–2)

Self-conceit can thus be occasioned by seeming acts of *beneficence*. Specifically, one falls into self-conceit when one mistakes the fulfilment of an obligatory duty for a meritorious act of beneficence. The failure to recognize one's existing obligations towards others occurs especially in the context of civil society. Thus, for example, in the discussion of moral reverence in the Doctrine of Method of the second *Critique*, Kant notes that

One need only reflect a little and one will always find a debt that he has somehow incurred with respect to the human race (even if it were only that, by the inequality of human beings in the civil constitution, one enjoys advantages on account of which others must all the more do without), which will prevent the self-complacent image of *merit* from supplanting the thought of *duty*.³⁸

We fool ourselves, in other words, when we ascribe merit to ourselves for an action that only attempts to set right the injustices generated by ‘the inequality of human beings in the civil constitution’.³⁹ And this, like any instance in which an agent ascribes a fictional virtue to herself, is an instance of self-conceit.

This last observation provides an important observation with which to conclude the discussion. Self-conceit, I have argued, should not be understood as the tendency or inclination to impose one’s unreasonable self-love upon others. It is at its core a *mode of thinking*, a failure of self-evaluation and self-knowledge. But self-conceit need not be motivated by an agent’s explicit desire to act on unreasonable self-love; it can also be brought about by an agent’s tendencies to mimic or imitate examples of virtue, rather than to apply critically the standards of practical reason. Self-conceit can thus often involve a kind of disturbing sincerity. The conceited agent does not reject the moral law, or even – like the psychopath – appeal to it in scare quotes or ‘as-if’ terms. Rather, she has an incomplete reverence for the law that does not extend to the critical and impartial scrutiny demanded of all agents as the ‘first command of all duties’ to themselves (1996a: 562; *MS*, 6: 441).

Self-conceit is, for this reason, an intractable and insidious problem. It is not, however, insoluble. Kant suggests that the voice of conscience is ever-present, even if faint. The agent who wishes to avoid self-conceit must engage in self-evaluation with ‘well-meant strictness’ in order to determine how far she falls short of ‘accordance with an uncompromising law’ (Kant 1996a: 263; *KpV*, 5: 154). Of course, if such a task were easy or straightforward, self-conceit would not pose the problem that it does. An agent who wishes to avoid self-conceit needs tools to help her avoid the errors of illusion and misinterpretation that lead easily to a false assessment of virtue. Toward this end, especially, we may be reminded of Kant’s remark regarding the formula of universal law, that ‘in moral judging it is always better to proceed by the strict method, and make the foundation the universal formula of the categorical imperative’ (Kant 2011a: 101; *G*, 4: 436). But further, agents should – as Kant urges – seek out moral relationships (especially friendships) with others who can point out to them the errors and oversights in their reasoning. Finding a reliable touchstone may in this sense be the surest way to avoid the cyclopic error of self-conceit.⁴⁰

Notes

- 1 Kant 1996a: 200; *KpV*, 5: 74. Abbreviations of Kant’s works are as follows: *Anth*: *Anthropology from a Pragmatic Point of View* (Kant 2008a); *EaD*: *End of All Things*

- (Kant 1996b); G: *Groundwork for the Metaphysics of Morals* (Kant 2011a); *KpV*: *Critique of Practical Reason* (Kant 1996a); *KrV*: *Critique of Pure Reason* (Kant 1998); *MS*: *Metaphysics of Morals* (Kant 1996c); *NEV*: Announcement of the Program of Kant's Lectures for the Winter Semester 1765–1766 (in Kant 2011b); *Päd*: *Lectures on Pedagogy* (Kant 2008c); *Refl*: *Reflections*; *RGV*: *Religion Within the Boundaries of Mere Reason* (Kant 1996d); *TG*: *Dreams of a Spirit Seer* (in Kant 2003); *VATP*: Draft for Theory and Practice; *VAMS*: *Drafts for the Metaphysics of Morals*; *VAnth*: *Lectures on Anthropology*; *VE*: *Lectures on Ethics* (Kant 1997); *VGSE*: *Drafts for Observations on the Feeling of the Beautiful and Sublime*; *VK*: *Essay on the Maladies of the Head* (Kant 2008b); *VL*: *Lectures on Logic* (Kant 1992); *ZeF*: *Toward Perpetual Peace* (Kant 1996e).
- 2 A notable exception is Robin Dillon, whose interpretation of one type of arrogance comes close to my analysis of self-conceit. I discuss Dillon's account in §4.
 - 3 Darwall 2006: 135–6. On the account that Beck, Reath and Darwall advance, the conceited agent need not think of others as *worthless*; he merely thinks of himself as having a special status with respect to moral legislation. This interpretation thus contains an important observation: the conceited agent's failures are failures of self-regard. However, the tendency to impose principles upon others and 'dominate' others is, I argue, a symptom of a deeper and more insidious failure of self-knowledge and self-evaluation.
 - 4 Engstrom's interpretation comes closest to the view advanced in this paper, since he emphasizes the extent to which moral self-conceit depends on an awareness (and possibly acceptance) of the moral law. Compare Ameriks (2012: 164) who argues that self-conceit requires that the agent accept the moral law's authority, and Sussman (2008) who argues that acquaintance with the moral law is sufficient.
 - 5 Hill (1998: 267). Hill suggests that this points to the unity of reason. But see Kleingeld (1998: 316–17), who argues that Kant's claims about the unity of reason cannot be explained by the fact that theoretical and practical reasoning can fail in the same ways.
 - 6 Kant 1996a: 201; *KpV*, 5: 75; Kant 2008a: 310; *Anth*, 7: 203.
 - 7 Examples follow in the discussion below. See also Kant's claim in the Doctrine of Method of the first *Critique* that 'it is not suited to the nature of philosophy, especially in the field of pure reason, to strut about with a dogmatic gait and to decorate itself with the titles and ribbons of mathematics. ... These are idle pretensions that can never succeed, but that instead countermand its aim of revealing the deceptions of a reason that misjudges its own boundaries and of bringing the self-conceit of speculation back to modest but thorough self-knowledge by means of a sufficient illumination of our concepts' (Kant 1998: 641; *KrV*, A735/B763).
 - 8 Kant 2011b: 251; *NEV* 2: 305. Compare Kant's remark in the lectures on pedagogy: 'Nothing is more ridiculous than precocious modesty or cheeky presumption (*naseweiser Eigendünkel*) in a child' (Kant 2008c: 459; *Päd*, 9: 469).
 - 9 In *Perpetual Peace* Kant observes that 'it is absurd to conclude from a single event to a particular principle of the efficient cause (to conclude that this event is an end and not merely an indirect result, by a natural mechanism, of another end quite unknown to us), and it is full of self-conceit (*Eigendünkel*), however pious and humble such talk may sound' (Kant 1996c: 331; *ZeF*, 8: 361n.).
 - 10 For references to self-conceit in philosophical debate, see *TG* 2: 353, *VL* 24: 151, *VL* 24: 873–5. For scientific debate, see *Refl*. 15: 394.
 - 11 For detailed and interesting discussion of errors in reasoning, see Piper's (2013: 289–316) discussion of 'pseudorationality'. Given the threat of self-conceit, it is unsurprising that Kant suggests a particularly strong antidote to the problem – scepticism. Kant not only

recommends scepticism as a way of overcoming self-conceit, but he also suggests that this is the only legitimate use of scepticism (*KrV*, A757/B785). But excessive scepticism can become its own form of self-conceit. This occurs when a person's scepticism begins to serve as an argument for a positive claim. Such a claim would be as presumptuous as its original target. Kant thus warns that one should abandon scepticism 'as soon as he has finished off the dogmatic self-conceit (*Eigendünkel*) of his opponent', since 'as soon as he would make [his] objections valid as a proof of the opposite his claim would be no less proud and conceited (*stolz und eingebildet*) than if he had seized hold of the affirmative part and its assertion' (Kant 1998: 664; *KrV*, A781/B809).

- 12 For example: *MS*, 5: 86 and *RGV*, 6:173. For clarity's sake, I depart from Gregor's translation of *Schwärmerei* as 'enthusiasm' and instead use Werner Pluhar's 'fanaticism'. This choice of terminology better preserves Kant's own distinction between *Schwärmerei* and *Enthusiasmus*. *Schwärmerei*, according to Kant, resembles moral feeling, but does not take the moral law as its object. It is akin to a passion that replaces the representation of duty with a type of self-congratulatory sentimentalism. *Enthusiasmus* appears to be more akin to an affect (see *MS* 6: 409) that takes instances of morally commendable behaviour as its object. For example, Kant thinks there can be enthusiasts of patriotism and friendship. For Kant's remarks on *Enthusiasmus*, see *Anth*, 7: 267 and *VAnth*, 25: 1006. Kant appears to be far less critical of *Enthusiasmus* than he is of *Schwärmerei*, though the former is still no substitute for the representation of duty.
- 13 This feature of self-conceit is apparent in comparisons that Kant makes between his theory and those of the Stoics and Epicureans, especially with respect to ideas of moral progress and the highest good. Kant remarks that 'The *summum bonum* of the philosophical sects could only take place when one accepted that humans could be adequate to the moral law. To that end, one had to either favorably construe his actions with moral self-conceit (*Eigendünkel*) or make the moral law very lenient. [However], the Christian can know the frailty of his personal worth and still hope himself to participate in the highest good under the condition of the holiest laws' (*Refl*, 19: 187). The Epicurean dilutes the demands of morality, such that it is difficult to fail to abide by its guidelines. The Stoic, on the other hand, maintains a stringent moral code, but allows for a type of self-conceit when it comes to an agent's evaluation of himself. As Kant remarks in another fragment, the Stoic tends to 'trust in his strength in the face of all temptation' (*Refl*, 19: 309). The virtuous Kantian agent recognizes the demandingness of morality, and simultaneously acknowledges that his is a continuous struggle to live up to those demands.
- 14 The observation appears in the early (pre-autonomy) Kaehler/Collins lectures (*VE-Collins*, 27: 350), and in the later Vigilantius lectures (*VE-Vigilantius*, 27: 611).
- 15 The Blomberg lectures on logic identify 'the principium of self-conceit' (*Eigen-dünkels*) with the tendency to think that the 'judgements of others are ... utterly dispensable in the use of [one's] own reason and for the cognition of truth' (Kant 1992: 119; *VL-Blomberg*, 24: 151).
- 16 Following Chignell (2007), I take it that Kant is operating with two senses of 'subjective sufficiency' in the discussion in the Canon of Pure Reason. The first (and the one at issue when it comes to *Überredung*) is the basis of mere opinion when it is not coupled with objective sufficiency. The second sense of 'subjective sufficiency' is the type of justification associated with practical belief.
- 17 Similar observations about communicability appear in the lectures on logic in the context of logical egoism. Kant defines logical egoism as the 'prejudice in accordance with which we hold the agreement of our understanding with the reason of others to be unnecessary as a *critierium* of wisdom' (Kant 1992: 323; *VL-Vienna* 24: 873). Elsewhere,

- Kant describes logical egoism as a prejudice in which ‘one believes ... and fully persuades oneself (*sich völlig überredet*), that one simply does not need the help of the judgements of others in a judgement of the understanding’ (Kant 1992: 141; *VL-Blomberg*, 24: 178). And, of course, in the practical sphere, Kant remarks that ‘[a]ll actions relating to the rights of others are wrong if their maxim is incompatible with a principle of publicity’ (Kant 1996e: 347; *ZeF*, 8: 381).
- 18 In the *Anthropology*, Kant claims that self-conceit and arrogance border on dementia (*Wahnsinn*) (*Anth*, 7: 203). Similarly, in a fragment on anthropology, Kant observes that self-conceit borders on derangement (*Verrückung*) (*Refl*, 15: 209).
 - 19 ‘Logical egoism is either *indifference* (*Gleichgültigkeit*) toward the judgements of others ... or it is conceit (*Eigendünkel*) and arrogance where one allots it to himself to make a correct judgement for all others’ (Kant 1992: 323; *VL-Vienna*, 24: 873). Two fragments on logic make the same distinction between indifference (*Gleichgültigkeit*) and self-conceit (*Eigendünkel*). See *Refl*. 16: 418 and 424.
 - 20 That self-conceit (*Eigendünkel*) is a mode of thought may find further support in the term’s etymology: though ‘-*dünkel*’ may sound similar to ‘*dunkel*’ (darkness), Grimm locates a closer connection to ‘*dünken*’ or ‘*denken*’, i.e. ‘thinking’ (Grimm Dictionary Online). So e.g. Martin Luther claims that it is either the devil or one’s own thoughts (*ihr eigen Dünkel*), which the devil controls, that causes people to worship false deities (Luther 1883–2009, 19: 207.19).
 - 21 The Collins lecture on moral philosophy, though dated to the winter semester of 1784/1785, is a copy of the lecture notes transcribed by Kaehler a decade or so earlier. Quotations here reflect the Cambridge translation of the Collins lectures.
 - 22 Excessive well-liking (*Wohlgefallen*) of oneself and self-conceit are identified explicitly at *KpV*, 5: 73. Kant makes a similar association at *VE-Vigilantius*, 27: 621, and *RGV*, 6: 45n., though in neither of those passages does he use the term *Eigendünkel*. However, in both of those passages self-love is associated with well-wishing (*Wohll wollen*) towards oneself, and in both passages well-liking (*Wohlgefallen*) of oneself is associated with the esteem one has for oneself morally.
 - 23 Note, too, that Kant’s agreement with Baumgarten’s perfectionist gloss is consistent with the date of these lectures, since they precede the *Groundwork*’s turn to autonomy by about a decade.
 - 24 *VE-Vigilantius*, 27: 621. Overlooking this point can lead to unnecessary confusion: Lewis White Beck notes in his commentary on the second *Critique* that Kant’s notion of self-love appears to vary from text to text. According to Beck’s assessment, there are moments at which Kant’s understanding of self-love captures an ‘egoism of feelings’ or an excessive concern that one’s inclinations be fulfilled; this understanding of self-love is especially apparent in the distinction between self-love and self-conceit in the second *Critique*, and in Kant’s description of self-love as a kind of self-regarding well-wishing. At other moments, as in this passage from the Kaehler/Collins lectures, self-love captures a broader notion of ‘egoism of the will’, something that comes much closer to self-conceit (Beck 1960: 100 n. 20). But as we have already seen, the early statement of self-love from the Kaehler/Collins lecture is a reference to Baumgarten, and Kant’s later discussions of self-love preserve (perhaps at the price of clarity) Baumgarten’s umbrella notion of *philautia*. An exception is the second *Critique*, where Kant uses the term *Selbstsucht* (self-regard) as the umbrella term instead.
 - 25 In the *Religion*, Kant implies that there are reasonable and unreasonable forms of each, since he notes that both types of love ‘must be rational’ (i.e. if either form of love could not be constrained by reason to be consistent with reason, this proviso would be unnecessary).

- 26 In the *Vigilantius* lecture, Kant makes the connection with ends explicit: ‘Love, generally speaking, is opposite to the will’s determination to strict duty and consists in the inclination or will to promote the ends of others’ (Kant 1997: 363; *VE-Vigilantius*, 27: 620).
- 27 Kant 1997: 363; *VE-Vigilantius*, 27: 261. It is worth noting the similarity of this claim to Kant’s claim in the second *Critique* that ‘reasonable self-love’ (*vernünftige Selbstliebe*) is the self-love that we constrain to the ‘condition of agreement’ with the moral law (Kant 1996a: 199; *KpV* 5: 73). There is a subtle difference between the two accounts: according to the account offered in the second *Critique*, self-love is permissible if it is consistent with the moral law. According to the *Vigilantius* lecture, self-love is permissible if it is not solipsistic, i.e. if we also have a love of others, and adopt their ends as our own. Of course, the universality requirement embedded in the moral law ensures that these two claims are at least extensionally equivalent.
- 28 ‘*The love of well-liking toward oneself*. This, too, is *philautia*, if it is exclusively entertained toward oneself, but also becomes unreasonable’ (Kant 1997: 364; *VE-Vigilantius* 27: 621).
- 29 For example, *KpV* 5: 154; *VE-Vigilantius* 27: 609–10.
- 30 See Dillon (2004: 204). Dillon argues further that ‘primary arrogance’ or self-conceit need not give rise to what she calls ‘interpersonal arrogance’, or the denial of another’s dignity and self-respect (191–6). In the paragraphs that follow, I disagree with this claim to some extent (see n. 33).
- 31 This point is not merely a claim about unreasonable self-love on the Kantian account. It would seem to be a feature of moral failing on any account that takes impartiality seriously. Henry Sidgwick makes a similar point regarding a utilitarian theory: ‘When, however, the Egoist puts forward, implicitly or explicitly, the proposition that his happiness or pleasure is Good, not only *for him* but from the point of view of the Universe, – as (e.g.) by saying that “nature designed him to seek his own happiness,” – it then becomes relevant to point out that *his* happiness cannot be a more important part of the Good, taken universally, than the equal happiness of any other person’ (1981: 420–1).
- 32 See Sensen (2009). Compare Korsgaard (1996: 123), who interprets claims like this as claims about a value that all rational agents share: ‘The unconditional condition of the goodness of anything is rational nature, or the power of rational choice. To play this role, however, rational nature must itself be something of unconditional value – an end in itself.’
- 33 Here I disagree, if in a merely Pickwickian sense, with Dillon’s claim that ‘primary arrogance’ (i.e. self-conceit) need not imply the denial of another person’s dignity. If Kant’s notion of dignity is indeed a relational notion, rather than a claim about a value property, then to evince self-conceit is just to deny that one stands in a particular relation to other agents, i.e. it is to deny their dignity. Still, I agree with Dillon’s core claim, namely, that self-conceit need not imply or require any judgement about the value or *worthlessness* of others.
- 34 *VAMS*, 23: 406, author’s emphasis.
- 35 Here I follow Timmermann’s (2006: 508–15) interpretation of Kant’s ‘primacy’ claim as being the claim that any observation of duty (from the motive of duty) necessarily implies ‘internal’ autonomous willing.
- 36 For a detailed discussion of conscience in the *Doctrine of Virtue*, see Esser (2013).
- 37 *Ibid.* Here, Gregor translates *schwärmerische* as ‘fanatical’, agreeing with Pluhar.
- 38 *KpV*, 5: 155n. For similar claims, see also *VE-Collins*, 27: 416 and *MS*, 6: 454.
- 39 Material inequality, as such, is not an injustice on Kant’s view. Injustice concerns ‘only a [person’s] relation to [another’s] *choice*’ (Kant 1996c: 387; *MS*, 230). When one

person's innate right to freedom is coopted by another's power of choice, this is injustice. Material inequality can, however, easily lead to injustice in civil society if it brings about this scenario. The person e.g. who relies on the assistance of others for survival in civil society suffers from an injustice; his innate right to freedom is placed at the whim of those who would offer him assistance (see Ripstein 2009: 274). Kant's point is just that when a person with means (who presumably also benefits from this injustice) gives the destitute person money, she is rectifying an injustice, not being beneficent. She should, therefore, not ascribe to herself any the merit associated with being beneficent.

40 I am grateful to Mavis Biss, Yoon Choi, Jeppe von Platz, Martin Sticker, Jens Timmermann, and two anonymous reviewers for this journal for their comments on earlier versions of this paper. I would also like to thank the audience at the workshop on self-knowledge, self-conceit and self-deception at Suffolk University for a helpful discussion of the paper. Much of the research for this paper was made possible by a Theodore and Jane Norman Junior Faculty Research Leave, and I am grateful to Brandeis University and the Norman family for this fellowship.

References

- Ameriks, K. (2013) 'Is Practical Justification in Kant Ultimately Dogmatic?'. In M. Timmons and S. Baiasu (eds), *Kant on Practical Justification* (Oxford: Oxford University Press), 153–75.
- Baumgarten, A. (1969) *Ethica Philosophica*. Hildesheim: G. Olms.
- Beck, L. W. (1960) *A Commentary on Kant's Critique of Practical Reason*. Chicago: University of Chicago Press.
- Chignell, A. (2007) 'Belief in Kant'. *Philosophical Review*, 116, 323–60.
- Darwall, S. (2006) *The Second-Person Standpoint: Morality, Respect, and Accountability*. Cambridge, MA: Harvard University Press.
- Dillon, R. (2004) 'Kant on Arrogance and Self-Respect'. In C. Calhoun (ed.), *Setting the Moral Compass: Essays by Women Philosophers* (Oxford: Oxford University Press), 191–216.
- Engstrom, S. (2010) 'The *Triebfeder* of Pure Practical Reason'. In A. Reath and J. Timmermann (eds), *Kant's Critique of Practical Reason: A Critical Guide* (Cambridge: Cambridge University Press), 90–118.
- Esser, A. M. (2013) 'The Inner Court of Conscience, Moral Self-Knowledge, and the Proper Object of Duty'. In A. Trampota, O. Sensen and J. Timmermann (eds), *Kant's Tugendlehre: A Comprehensive Commentary* (Berlin: De Gruyter), 269–92.
- Hill, T. E. (1998) 'Kant's Argument for the Rationality of Moral Conduct'. In P. Guyer (ed.), *Kant's Groundwork of the Metaphysics of Morals: Critical Essays* (New York: Rowman & Littlefield), 249–72.
- Kant, I. (1902–) *Kants Gesammelte Schriften*. Berlin: de Gruyter.
- (1992) *Lectures on Logic*, trans. and ed. J. M. Young. Cambridge: Cambridge University Press.
- (1996a) *Critique of Practical Reason*, trans. M. Gregor. In M. Gregor (ed.), *Practical Philosophy*. Cambridge: Cambridge University Press.
- (1996b) *The End of All Things*, trans. A. W. Wood.. In A. W. Wood and G. di Giovanni (eds), *Religion and Rational Theology*. Cambridge: Cambridge University Press.
- (1996c) *The Metaphysics of Morals*, trans. M. Gregor. In M. Gregor (ed.), *Practical Philosophy*. Cambridge: Cambridge University Press.
- (1996d) *Religion Within the Boundaries of Mere Reason*, trans. G. di Giovanni. In A. W. Wood and G. di Giovanni (eds), *Religion and Rational Theology*. Cambridge: Cambridge University Press.

- (1996e) *Toward Perpetual Peace*, trans. M. Gregor. In M. Gregor (ed.), *Practical Philosophy*. Cambridge: Cambridge University Press.
- (1997) *Lectures on Ethics*, trans. P. Heath. ed. P. Heath and J. B. Schneewind. Cambridge: Cambridge University Press.
- (1998) *Critique of Pure Reason*, trans. and ed. P. Guyer and A. W. Wood. Cambridge: Cambridge University Press.
- (2003) *Theoretical Philosophy 1755–1770*, trans. and ed. D. Walford and R. Meerbote. Cambridge: Cambridge University Press.
- (2008a) *Anthropology from a Pragmatic Point of View*, trans. R. Loudon. In R. Loudon and G. Zöllner (eds), *Anthropology, History, and Education*. Cambridge: Cambridge University Press.
- (2008b) *Essay on the Maladies of the Head*, trans. H. Wilson. In R. Loudon and G. Zöllner (eds), *Anthropology, History, and Education*. Cambridge: Cambridge University Press.
- (2008c) *Lectures on Pedagogy*, trans. R. Loudon. In R. Loudon and G. Zöllner (eds), *Anthropology, History, and Education*. Cambridge: Cambridge University Press.
- (2011a) *Groundwork of the Metaphysics of Morals*, trans. M. Gregor and J. Timmermann. Cambridge: Cambridge University Press.
- (2011b) *Observations on the Feeling of the Beautiful and Sublime and Other Writings*, trans. and ed. P. Frierson and P. Guyer. Cambridge: Cambridge University Press.
- Kleingeld, P. (1998) 'Kant on the Unity of Theoretical and Practical Reason'. *Review of Metaphysics*, 2, 311–39.
- Korsgaard, C. (1996) *Creating the Kingdom of Ends*. New York: Cambridge University Press.
- Luther, M. (1883–2009) *Martin Luthers Werke: Kritische Gesamtausgabe*. Weimar: Hermann Böhlau.
- Piper, A. (2013) *Rationality and the Structure of the Self*. vol. 2: *A Kantian Conception*. Berlin: APRA Foundation.
- Reath, A. (2006) 'Kant's Theory of Moral Sensibility'. In *Agency and Autonomy in Kant's Moral Theory* (Oxford: Oxford University Press), 8–32.
- Ripstein, A. (2009) *Force and Freedom: Kant's Legal and Political Philosophy*. Cambridge, MA: Harvard University Press.
- Sensen, O. (2009) 'Kant's Conception of Human Dignity'. *Kant-Studien*, 100, 309–31.
- Sidgwick, H. (1981) *The Methods of Ethics*. Indianapolis: Hackett.
- Sussman, D. (2008) 'From Deduction to Deed: Kant's Grounding of the Moral Law'. *Kantian Review*, 13, 52–81.
- Timmermann, J. (2006) 'Kantian Duties to the Self, Explained and Defended'. *Philosophy*, 81, 505–31.
- Wood, A. (1996) 'Self-Love, Self-Benevolence, and Self-Conceit'. In S. Engstrom and J. Whiting (eds), *Aristotle, Kant, and the Stoics* (Cambridge: Cambridge University Press), 141–61.