

A Funny Thing Happened on the Way to Self-Knowledge

Jordan MacKenzie

Forthcoming at *The Monist*: Please cite published version when available

Abstract

What moral value, if any, is there to unsuccessful self-scrutiny? I answer this question by drawing on Kant's discussion of the duty of self-knowledge in the *Metaphysics of Morals*. On my interpretation, the duty of self-knowledge is a demand to understand the generic structure of our moral agency: fulfilling it requires that we appreciate ourselves as fundamentally flawed, but also fundamentally worthy moral agents. This interpretation has two exegetical upshots: it renders the duty consistent with Kant's skepticism about the possibility of moral self-knowledge and explains why Kant called self-knowledge the 'First Command' of all self-regarding duties. But this interpretation has value beyond the exegetical. Specifically, by appreciating the value of generic moral self-knowledge, we can understand why self-scrutiny can be worthwhile even when it does not yield particular self-knowledge.

Introduction

It finally happened. You did the thing that you promised yourself you'd never do. Perhaps you cheated on your spouse, or stole from your grandparents, or abandoned your best friend in their hour of need.

But . . . *why* did you do it? Was your unforgivable act the result of a simple lapse in judgment, or a good intention gone awry? Were you spurred on by repressed sexual jealousy, or

unresolved issues from your childhood? Or are you perhaps mistaken in thinking that you've actually done something wrong? Perhaps your feelings of guilt are just holdovers from your strict religious upbringing or internalized self-loathing.

Self-knowledge is not always easy to come by. No matter how much we scrutinize ourselves, we may still wonder whether we've actually gotten a handle on who we are. And the reasons that we often fail to know ourselves are both vexing and commonplace: we trade uncomfortable truths for pleasant fantasies, we mistake patterns of behavior for momentary lapses in judgment, and we give ourselves the benefit of the doubt even in situations wherein we would not let others off the hook. Even if these psychological shortcomings don't entirely rule out the possibility of knowing our motives and characters, we must grant that they often make the task very difficult.

And yet, we persevere. We continue to pay our therapists and write in our diaries. The pursuit of self-knowledge, further, can feel not only prudentially valuable but morally mandatory. A lack of self-knowledge, after all, is not one epistemic failing among many. It's a moral shortcoming (Cholbi 2022; MacKenzie 2018).

Let's grant for the sake of argument that the psychological picture I've sketched out is plausible: we are, in other words, self-opaque. If this is right, then we might wonder what exactly we get from committing ourselves to self-knowledge—especially when that commitment often doesn't bear fruit. What, if anything, is morally valuable about unsuccessful self-scrutiny?

In this paper, I use Kant's discussion of the duty of self-knowledge in the *Metaphysics of Morals* to answer this question. My paper is, as such, largely exegetical. But I think that this exegesis has a payoff for contemporary discussions of self-knowledge and self-regarding

obligations. Specifically, we can use Kant's text to explain why self-scrutiny can be morally valuable *even if* it is unlikely to give us knowledge of our particular characters or motives.

My paper proceeds as follows. Section 1 outlines the content of Kant's duty of self-knowledge and introduces two puzzles: the first has to do with the exegetical consistency of the duty given Kant's broader commitment to human self-opacity, while the second questions the cogency of its status as the 'First Command' of all duties. Section 2 draws on work by Jeanine Grenberg to argue that we can escape both puzzles by interpreting the duty of self-knowledge as a duty that primarily helps us achieve an awareness of our generic moral worth. While the duty commands us to scrutinize our characters and motives, it does so with the aim of having us fix our eyes on our fundamental moral goodness and inescapable moral imperfection. I wrap up in Section 3 by extending this story to contemporary questions about the moral value of self-knowledge and self-scrutiny. Even unsuccessful self-scrutiny helps us to acknowledge a conception of our moral agency that we need to have in place to avoid common moral failings. The very act of scrutiny, further, is a means by which we affirm ourselves to have the sort of worth that makes coherent self- and other-regarding obligations.

1. Kant's First Command

In the Doctrine of Virtue in the *Metaphysics of Morals*, Kant makes a surprising claim. The 'First Command' of all self-regarding duties, he contends, is to:

...*know* (scrutinize, fathom) *yourself*, not in terms of your natural perfection (your fitness or unfitness for all sorts of discretionary or even commanded ends) but rather in terms of your moral perfection in relation to your duty. That is, know your heart—whether it is good or evil, whether the source of your actions is pure or impure, and what can be

imputed to you as belonging originally to the *substance* of a human being or as derived (acquired or developed) and belonging to your moral *condition*. (MM 6:441)¹

I'll break this claim down a little, before explaining why it's surprising.

Let's start by clarifying the content of the duty. Kant's 'First Command' actually demands the pursuit of two different types of knowledge. First, the duty commands us to know our hearts (our settled moral disposition to prioritize either self-love or the moral law; MM 6:441; R 6:30), which entails knowing whether we are good or evil, and whether our actions spring from morally pure or impure sources. Call this sort of self-knowledge *particular* self-knowledge,² as it is knowledge of our individual moral characters.

Second, this duty commands that we seek to acquire what I will call *generic* self-knowledge, which is knowledge related to our human 'substance'. For Kant, human nature is fundamentally dualistic. We are, on the one hand, moral beings capable of willing the moral law and thus acting from duty. The moral law, further, is something we experience as the supreme law of our practical reason and thus as what gives us moral worth (MM 6:441; CPrR 5:73, 5:77,

1. All English translations are from *The Cambridge Edition of the Works of Immanuel Kant*, with the exception of the *Critique of Pure Reason*, which is the Norman Kemp Smith translation (Palgrave MacMillian). Kant's titles are abbreviated as follows. *A*=*Anthropology from a Pragmatic Point of View*; *G*=*Groundwork for the Metaphysics of Morals*; *CPR*=*Critique of Pure Reason*; *CPrR*=*Critique of Practical Reason*; *MM*=*Metaphysics of Morals*; *R*=*Religion within the Boundaries of Mere Reason*; *C*=*Collins Lectures on Ethics*; *H*=*Herder Lectures on Ethics*, *V*=*Vigilantius Lectures on Ethics*.

2. The terms 'particular' and 'generic' self-knowledge come from Ware (2009).

5:87–89; *R* 6:23). But in addition to having this fundamental goodness, we also have within us a “propensity to evil” (*R* 6:31–32). By ‘evil’, Kant has in mind our tendency to deviate from the moral law while remaining conscious of its authority (*R* 6:32). To recognize the substance of our humanity, then, is to both recognize that we are fundamentally morally good and fundamentally morally imperfect.

In addition to telling us that the duty requires particular and generic self-knowledge, Kant also tells us what we will achieve by fulfilling it:

Moral cognition of oneself, which seeks to penetrate into the depths (the abyss) of one’s heart which are quite difficult to fathom, is the beginning of all human wisdom. For in the case of a human being, the ultimate wisdom, which consists in the harmony of a human being’s will with its final end, requires him first to remove the obstacle within (an evil will actually present in him) and then to develop the original predisposition to a good will within him, which can never be lost. (Only the descent into the hell of self-cognition can pave the way to godliness.) (*MM* 6:441)

This moral self-cognition, Kant contends, will dispel any contempt that we have for ourselves as human beings, and it will also “counteract that *egotistical* self-esteem which takes mere wishes—wishes that, however ardent, always remain empty of good deeds—for proof of a good heart” (6:441–42). Through moral self-cognition, in other words, we can avoid the dual vices of timorousness and self-conceit, which Kant describes elsewhere as the “two rocks a man runs into if he departs in one direction or the other, from the moral law” (*C* 27:351).

Now for the surprise—or rather, surprises. The first reason that the duty of self-knowledge is surprising has to do with its apparent *exegetical inconsistency*. Specifically, the duty seems boldly inconsistent with Kant’s well-established stance on human self-opacity. We

have, as Kant repeatedly stressed, a deep-seated tendency to self-conceitedly overestimate the goodness of our actions, the quality of our hearts, and the source of our moral worth (*MM* 6:447, 6:393; *G* 4:407; *CPR* A551/B579; *R* 6:51). Call this puzzle the *Consistency Puzzle*.

But even if we put these exegetical worries aside, it's also surprising that *self-knowledge* would be the 'First Command' of all self-regarding, and thus other-regarding,³ duties. Is Kant telling us that self-knowledge is a prerequisite for moral agency? Or is he implying that all of our other duties are derivable from the duty of self-knowledge? Call this puzzle the *Priority Puzzle*.

Previous scholarship has focused on the Consistency Puzzle.⁴ Resolving the consistency question requires one to find some form of self-knowledge that isn't ruled out by Kant's claims about self-opacity or to find some method of self-scrutiny that can evade the distorting force of self-conceit. But a successful resolution to the Consistency Question does not, by itself, suggest any particular answer to the Priority Puzzle. It's one thing to figure out how to know ourselves; it's another thing to figure out why Kant assigned self-knowledge the status of a 'First Command'.

I'm going to suggest that we gain something by trying to tackle the Priority Puzzle first. Specifically, I think that a successful resolution to that puzzle suggests a particular story about what sort of self-knowledge is most central to Kant's duty of self-knowledge, and about how the duty can be rendered compatible with his broader thesis on human self-opacity. It is to this matter that I now turn.

3. *C* 27:314; *MM* 6:417–18; see also Denis (2012).

4. See, e.g., O'Hagan (2009), O'Neill (1996), and Ware (2009).

2. The Priority and Consistency of Kant's First Command

What could Kant possibly mean by the claim that self-knowledge is the 'First Command' of all self- (and thus other-) regarding duties?

Perhaps Kant is making some sort of *ontological* point. Some duties could be said to 'derive' from more basic moral obligations. Indeed, when discussing the duty of self-knowledge, Kant identifies two such derivative duties: "impartiality in appraising oneself in comparison with the law, and sincerity in acknowledging to oneself one's inner moral worth or lack of worth" (*MM*: 6:441–42). Here, Kant seems to be pointing out that, once we see ourselves as under a duty to know ourselves, we'll also see ourselves as obligated to take the means to achieve that end. In this way, we can say that the existence of the primary duty grounds the existence of these derivative duties: if we were not obligated to know ourselves, we would also not be obligated to make efforts to know ourselves. But it is hard to see how we can tell a similar story about all other duties. Even if we turned out to be hopelessly self-opaque, we could still presumably have some *other* moral obligations.

I think it's more likely that Kant was making a point about *practical priority*: self-knowledge is something that we must have to fulfill other duties. But why would we need to know ourselves before we could go about fulfilling our other moral obligations? And what do we need to know?

We can make headway towards answering both of these questions by drawing an analogy with another one of Kant's epistemic 'firsts'. In the *Critique of Pure Reason*, Kant identifies the principle of the synthetic unity of apperception as the first principle of transcendental philosophy. The very possibility of cognizing objects, according to this principle, requires that all representations of those objects be subject to the unity of consciousness (*CPR* A117n, B138–39,

B407). For my thoughts to be contentful, in other words, I must be able to represent them (i.e., to affix the ‘I think’ to their representation) as belonging to a single consciousness.

There are two takeaways from the transcendental unity of apperception (‘TUA’) that are relevant to our present inquiry. The first is that, for there to be a thing that is recognizable as an ‘I’, it must be possible for a set of disparate experiences to be unified within a single locus of consciousness. The second is that the TUA gives us a form of self-knowledge in the broadest sense of the term. To know that I am a single recognizable consciousness is to know *something* about myself.

In understanding the TUA as a form of self-knowledge, we should be careful not to overstate our case. The ‘I think’ tells us there is a self, but it does not give us knowledge *of* that self (*CPrR* A350; see also B135, B155, B158). Nevertheless, the knowledge that there is a self is indeed self-knowledge. This is a point that Kant actually emphasizes in his *Vigilantius* discussion of the duty of self-knowledge: “Just as in the metaphysical sense, self-knowledge is presupposed in apperception of the determinations present in us,” he observes, “so too is it presupposed in the moral sense” (*V* 27:608).⁵

For now, I’ll put aside the question of what type of self-knowledge is presupposed as a pre-requisite for moral agency. Instead, I’ll consider what sort of moral unity could play a role analogous to the role played by the TUA in Kant’s theoretical philosophy. The hallmark of a moral action is that it is describable by a maxim that is a categorical, rather than hypothetical,

5. Kant also draws an analogy between self-conceit (understood as moral fanaticism) and the fanaticism that occurs when we overstep the boundaries of theoretical reason (*CPrR* 5:85-5:86; Moran 2014, 424).

imperative. When we will this maxim, in other words, we will a maxim that would be a moral law for all similarly situated moral agents. The moral law thus provides us with *interpersonal* unity—it is by virtue of being able to will the moral law that we are potentially able to form a ‘Kingdom of Ends’ with other rational agents.

But the moral law also unifies us *intrapersonally*, and it is this form of unity that is most relevant to my present purposes. There is a type of intrapersonal unity, I will argue, that we must have in order to represent ourselves as bound by moral obligations across time.

To see what I mean by ‘moral unity’ here, consider someone who doesn’t have it: the self-conceited person. Kant views self-conceit, that is, a tendency to overestimate one’s moral worth, as both a deep-seated psychological propensity and a serious threat to moral action. Crucially, self-conceit involves a failure of self-knowledge: it arises when an agent mistakenly takes some contingent feature of herself, like wealth or social status (*C* 27:465) to be proof of an elevated moral worth. In doing so, she misunderstands something about herself: that her moral worth has nothing to do with any contingent feature of her social circumstance. And by mislocating the source of her moral worth, the self-conceited agent opens the door to moral capriciousness. If her vaunted moral worth places her above the command of the moral law, then what’s to stop her from ‘tinkering’ with it until it suits her ‘inclinations and convenience’? (*C* 27:465; see also *CPrR* 5:73-4; *R* 6:29; *MM* 6:442).

Why think the self-conceited person lacks moral unity? To answer, just think about the sort of morally inconsistent people that we all have in our lives. I’m sure most of us have at least one person who is quick to condemn others for eating meat whenever she’s on a vegan kick, but who will merrily return to scarfing down ribs as soon as the meat craving hits her. Or think about the sort of people who view cheating as a mortal sin, but who find masterful ways to excuse their

own infidelities. These sorts of people are morally disunified in the way that Kant describes; they govern themselves by an ever-changing set of moral standards and seem to think that the rules that apply to others somehow don't apply to themselves.

We must not overstate how radical this disunity is. For Kant, self-conceit is not an irreparably fallen state. The self-conceited agent, for instance, still experiences moral emotions, and cares about acting morally—and this suggests, as Engstrom notes, the self-conceited person must operate under “some antecedent, if only implicit, consciousness of the moral law” (2010, 111; see also Moran 2014; Russell 2020). And so, there is a sense in which the self-conceited agent retains a thin sort of moral unity: it is always *possible* for her to recognize herself as governed by a single moral standard. What the self-conceited agent lacks is a type of ‘thick’ moral unity: she does not, as a matter of fact, *actually* recognize herself as governed by this standard.

This thick disunity has serious implications for the possibility of moral agency. The self-conceited agent who thinks she can tinker with the moral law at will might be able to perform morally appropriate actions. But so long as she continues to see herself as standing above morality, she won't be able to perform them *as* moral actions (Grenberg 2005, 226). Instead, she will continue to mistakenly treat categorical imperatives as merely hypothetical. This tells something about the importance of practical unity: absent the internal practical unity provided by a recognition of the moral law within, the best that we will be able to hope for is to be able to obey the letter but not the spirit of the law (*R* 6:31).

2.1 What Sort of Self-Knowledge Has Practical Priority?

If we assume that the duty of self-knowledge has *practical priority* over other duties, then we arrive at the following story about self-knowledge's value. Self-knowledge, if it truly can dispel fanatical self-contempt and strike down self-conceit, is valuable as a prerequisite for any genuinely moral action at all. This gives us a natural way to understand the duty's status as a 'First Command': it is the command that we must fulfill if we are to be able to obtain the sort of thick moral unity that makes the proper representation of duty (and thus genuine moral action) more than theoretically possibility.

And so too does this interpretation suggest a particular story about the type of self-knowledge we're ultimately meant to arrive at by fulfilling Kant's 'First Command'. Recall that the duty of self-knowledge is actually a duty of self-knowledges: we're meant to both know ourselves as particular moral agents, and as generic moral beings who are capable of willing the moral law out of duty, and not love. But surely both types of self-knowledge can't be a 'First Command'—something has to come second.

Most commentators have understood Kant's duty of self-knowledge as a duty to obtain particular self-knowledge.⁶ And it's obvious why they've done so: particular self-knowledge

6. See, e.g., Biss (2015), Cureton (2017), Denis (2001), Fairbanks (2000), Halper (2023), O'Hagan (2009), O'Neill (1996), Sticker (2019), Thomason (2013), Ware (2009), and Wood (1999). Ware and Halper both grant some value to generic self-knowledge in their accounts. Grenberg (2005) recognizes the foundational role that generic self-knowledge plays in Kant's moral framework, but still ultimately treats the duty as one aimed at particular self-knowledge. Messina (2013), in contrast, shares my view that generic self-knowledge is the duty's primary

seems to be front and center in Kant's articulation of the duty, and its moral value is obvious. It's hard to make progress towards moral self-perfection, for instance, if one doesn't know how one is presently falling short. But particular self-knowledge is a poor candidate for a 'First Command'. Its value, after all, is derivative: obtaining it helps us in our quest towards moral self-perfection.⁷ But this picture assumes that we, as agents, already acknowledge that we have a duty of moral self-perfection. We must, in other words, recognize at least one other duty before we can grasp the moral value of particular self-knowledge.

We do better by interpreting the duty as a duty to obtain 'generic' self-knowledge. Recall that the practical priority reading suggests that this duty is the one that we must fulfill in order to grasp ourselves as bound by a system of duties. And this is exactly what generic self-knowledge is meant to help us obtain. It is meant to 'strike down' the self-conceit that drives us to misapprehend the source of our moral worth allowing us to gaze at the moral law within (*CPrR* 5:73; see also 5:25, 5:77, 5:87–88). An awareness of the moral law within, as well as our relation to it, thus provides us with an awareness of ourselves as both imperfect and respect-worthy. And this, in turn, gives us thick moral unity: when we understand our substance, we're able to comprehend ourselves as perpetually bound by the same moral law.

command, albeit one that is required insofar as it is essential to the development of moral virtue and not to the possibility of moral action more generally.

7. This is the interpretation endorsed in most of the secondary literature. See Denis (2001), Fairbanks (2000), Grenberg (2012), O'Hagan (2009), O'Neill (1996), and Ware (2009).

Note that I am not the first interpreter to draw a connection between the apprehension of the moral law within and moral self-unity. Jeanine Grenberg (2018), most notably,⁸ has suggested that the feeling that we have of the moral law gives us a moral unity that is a practical parallel to the theoretical unity of apperception:

Through moral feeling, I thus gain access to my self-as-practical subject, clothing the bare bones of that theoretically accessed self-as-subject, the Transcendental Unity of Apperception, with an apperceptive cognition of merely practical content: I know myself

8. See also Halper (2023) and Messina (2013). Other scholars (e.g., Moyer 2006) have argued that *conscience* provides us with this practical apperception. To be sure, there is a clear relation between our phenomenological experience of the moral law, and our experience of conscience. Nevertheless, I think that it is the perception of the moral law (and our relation to it), and not the experience of our conscience, that fundamentally provides us with unity as practical agents. To see why, recall that Kant analogizes conscience to a court of law. Within this court, we both defend and prosecute our actions, until our inner judge (conscience) releases a verdict (*MM* 6:440). But for this ruling to bind us in a way that grants us practical unity, we must already recognize our inner judge as having the power to issue binding verdicts. This is something we regularly refuse to do; we often dismiss the call of our conscience, and tweak moral precepts to give us the verdict we desire. For conscience to perform the role Kant thinks it does, we need a moral awareness of its authority prior to going in front of its court.

as subject not only as the locus of all my thoughts but also as a being who is the source of all moral obligation. (43; see also Grenberg 2005)⁹

Where I depart from Grenberg is on the matter of whether we are under a duty to achieve this practical apperception. For Grenberg, the feeling of the moral law within is automatic (*CPrR* 5:31, 5:80; see also *CPrR* 5:25; *G* 4:454; *MM* 6:339–40) and thus not a potential subject of duty. But I think we shouldn't overstate this automaticity: while thin moral unity is automatic (in the sense that any rational agent can recognize the moral law within), thick moral unity is not. To achieve the latter sort of unity requires us to actually pay attention to that law, as well as our relation to it. Given that this is something that we may fail to do (as the pervasive phenomenon of self-conceit makes clear), it can be a proper subject of duty.

That fulfilling the 'First Command' serves as a corrective to our tendency towards self-conceited disunity raises another challenge to its status as a duty. Specifically, if fulfilling the

9. Grenberg's earlier (2005) position is closer to my own. There, she described the duty as one requiring (among other things) generic knowledge of our substance. When someone recognizes herself as the subject of her actions (which comes from grasping oneself as "a dependent and corrupt but capable and dignified agent" [224]), she

recognizes herself as the author of her acts; and implicit in this recognition of authorship is a recognition of all those qualities which constitute an obligated agent: her obligation and her capacity to fulfill it, the potentially conflicting ends of happiness and duty, and of the human tendency to place the former above the latter. To recognize herself substantially is thus to understand herself as a practical subject in general, the practical equivalent of understanding the TUA. (225)

duty is meant to make it possible for us to represent ourselves as bound by duties, we might wonder how it is that we could possibly represent the ‘First Command’ as a *duty* prior to fulfilling it. Here, the fact that we are capable of misperceiving, but not fully escaping, the moral law within us once again becomes relevant. A self-conceited agent might initially begin a project of self-understanding for bad motives. She might, for instance, think that it’s just a fun thing to do or that doing it will provide her with more evidence of her moral superiority. But through completing that project, she will find herself humbled by the law that made that project a duty. She will thus retrospectively recognize self-knowledge as something she ought to obtain, even if her lack of self-knowledge was initially what stood in the way of that recognition.

2.2 Generic Self-Knowledge and the Consistency Question

So generic self-knowledge, that is, knowledge of the relationship that we bear to the moral law, is a prerequisite for thick moral unity, which in turn is the sort of unity that we need to properly represent ourselves as bound by moral obligations across time. This answers the Priority Puzzle.

Now let’s turn our attention to the Consistency Puzzle. Recall that this puzzle arises because the duty of self-knowledge seems to be squarely in tension with Kant’s claims about moral self-opacity. We have, as he repeatedly stressed, a deep-seated, self-conceited tendency to overestimate our moral worth and the goodness of our actions. To be sure, it’s not obvious that this challenge is intractable. Ware (2009, 694), for instance, locates a means of achieving particular self-knowledge in our ability to be conscious of whether we’ve sincerely examined our actions in front of our inner court of consciousness, while Halper (2023, 190) argues that we are not self-opaque when it comes to our moral failure. Nevertheless, even if we grant that self-

conceit does not completely rule out the possibility of particular moral self-knowledge (as I think we should), we must recognize that it makes the task very difficult. And so, we might wonder how committed we should be to the project of knowing our particular selves, given that that project is likely to fail more often than it succeeds. Surely, there's some other morally valuable project that we could pursue that has a higher likelihood of success.

Does self-conceit similarly threaten generic self-knowledge? At times, Kant seems to think so. Recall that self-conceit can make us prone to mistakenly locating the source of our moral worth in contingent features of ourselves (*C* 27:457–58; see also *C* 27:465; *H* 27:41; *MM* 436–37; *R* 6:27). Nevertheless, it is precisely self-conceit that the sight of the moral law strikes down, by reminding us that the law is itself the only standard by which our morals are to be judged. (*CPrR* 5:73; see also *CPrR* 5:25, 5:77, 5:87–88; Papish 2018, 160–70). And since the moral law is a 'Fact of Reason', this 'striking down' will remain possible no matter how deep our self-conceit. So, while we may partially lose sight of our generic moral selves, we are never entirely generically self-opaque.

How exactly do we get the moral law to 'strike down' our self-conceit? Here's where particular self-knowledge becomes relevant. Recall that the duty, as written, seems to be primarily focused on the acquisition of particular, rather than generic, self-knowledge. And yet, I have argued that it is this latter sort of self-knowledge at which the duty ultimately aims. If I'm right, then we might wonder why Kant discusses particular self-knowledge at all.

Particular self-knowledge matters to the duty because we arrive at generic self-knowledge through the experience of our particular moral selves. When we feel a surge of awe upon finding an unrecognized moral strength within us or a pang of guilt upon having done something bad, we come to appreciate the power and limits of our moral agency. And when we

take a moment to inquire into those feelings, we affirm our worth and imperfections: our moral selves are worth taking seriously, but they're not fully transparent. In this regard, the duty truly is a 'practical analogue' to the TUA. It is, after all, by scrutinizing particular empirical claims about the world that we are able to turn our attention to the original (i.e., general) apperceptive unity of our consciousness. To grasp the changing of the seasons, for instance, is to recognize that the empirical representations that we have of a tree's changing leaves are representations that are being experienced by a single consciousness.¹⁰

And this is the very story that we get in the *Vigilantius Lectures*. There, Kant informs us that the self-knowledge upon which practical apperception depends "consists in examination of our past state, or comparison of our actions with their dutifulness insofar as we fulfill or transgress the same" (*V* 27:608). This claim initially seems odd. Kant begins by referring to a form of generic moral self-knowledge (the moral self-knowledge that is equivalent to the self-knowledge required for transcendental apperception), but then moves on to describe a method of self-inquiry that aims at *particular* self-knowledge (comparing our actions to the moral law). What explains this shift? Kant is telling us, I think, that to gain awareness of our status as moral agents, we must undertake the very sort of empirical self-scrutiny that is often insufficient for delivering particular self-knowledge.¹¹

10. I owe this point to G. Anthony Bruno.

11. Indeed, although space prevents me from expanding on this point, we might think that *generic* self-knowledge actually creates the ideal conditions from which we may successfully pursue particular self-inquiry. Appreciating the moral law within is a humbling experience, and

It is through an awareness, however imperfect, of our particular moral selves that we are afforded an opportunity to obtain generic moral self-knowledge. This helps explain why Kant immediately qualifies the command ‘to know’ with two different epistemic activities: scrutinizing and fathoming. The latter activities are focused on the pursuit, rather the possession, of knowledge. And it is through this pursuit—whether it results in factual beliefs about our particular moral characters or not—that we arrive at an appreciation of the moral law within. This appreciation, in turn, gives us both the humility and awe that we need to properly represent ourselves as bound by duty across time.

3. Self-Scrutiny Without Self-Knowledge

So far, this discussion has been an exegetical project. But I think that the value of Kant’s discussion of the duty of self-knowledge is not internal to Kant’s own theory. Specifically, I think that Kant has something important to tell us about the moral value of self-scrutiny (even when it’s unsuccessful). Self-scrutiny is instrumentally valuable insofar as it helps us grasp what sort of moral agents we are. And it can also be intrinsically valuable, insofar as it affirms our moral worth.

Recently, renewed attention has been paid to the moral dimensions of self-knowledge (see, e.g., Roessler 2014; MacKenzie 2018; Cholbi 2022, forthcoming). This trend picks up on what I take to be a common sentiment outside philosophy: that an absence of self-knowledge can be a serious moral defect. But self-knowledge is not always easy to come by. And so, we might

that humility might help ensure that we do not jump to conclusions or help ourselves to undeserved self-certainty.

wonder what value we ought to place in the project of trying to figure ourselves out, given that we so often fail at it. Couldn't our moral efforts be put towards a project that has a greater likelihood of success?

What Kant draws our attention to is the fact that even unsuccessful attempts at knowing our particular selves can yield morally valuable self-knowledge. When it comes to our moral failings, for instance, we may never know with whether to attribute them to *akrasia*, malice, or carelessness. But by looking closely at these failings, we at the very least affirm that they're worth taking seriously, and that they represent failures to hold ourselves accountable to the moral standards that govern us. This insight can be uplifting: the fact that you care to figure out why you did wrong speaks to the concern you have for doing right. And it can also be humbling: the fact that you must question your motivations reminds you that you're not morally perfect. The value of moral self-scrutiny, then, lies not just in its instrumental benefits, but in the picture of moral agency that it helps us to affirm.

We can motivate this story by thinking through an example. Imagine that you've just finished the grading for your last seminar, and you come to the realization that the top five marks in the class have all gone to white male students. This realization might be uncomfortable. You might find yourself wondering whether your grading standards were biased, or whether it was sheer coincidence that the best students in the class happened to be white and male.

It's possible that self-scrutiny won't lead you to self-knowledge. And yet, I think there's something valuable about trying to reckon with the question of whether your grading is biased in the absence of achieving an answer to it. For starters, the mere fact that you're trying to figure things out is itself some evidence that you care about doing right by your students. And so too does your self-scrutiny reflect a recognition of recognition of your own limits: you recognize that

your own desire to have the question answered in a self-flattering way doesn't itself mean that the question has been answered.

This picture of the self, as both morally worthy and morally flawed, is the picture we need to have if we are to hold ourselves morally accountable. And having it is not a trivial moral accomplishment. To go back to the previous example, I'm sure we've all encountered people who couldn't even *fathom* the possibility that their grading might reflect implicit biases. These people might react to the realization that the top marks in their class all went to white men with some modicum of guilt or discomfort, but they'll be quick to assure themselves that they have nothing to feel bad about. And they might reassure themselves by calling to mind times in which they exhibited a commitment to racial or gender justice. In any particular case, this sort of self-flatterer might be correct in their self-interpretation. Nevertheless, the tendency to view oneself as morally beyond reproach, or to take past actions as proof of an ongoing good will, does not generally lead to moral progress.

And so too do we all know people who are positively consumed by feelings of guilt and who view themselves as morally hopeless. Self-loathing, too, represents a failure of generic self-knowledge. The person who is utterly consumed by guilt, after all, is failing to realize what that guilt ultimately says about them. That they're capable of feeling guilt speaks to their commitment to morality: they wouldn't feel bad if they didn't care. Even if that commitment is never perfectly realized, it is still morally admirable. The fact that they care should be enough, in other words, to convince them that they aren't morally beyond repair.

So, unsuccessful self-scrutiny can be morally valuable insofar as it yields generic knowledge of our moral agency. And we need this knowledge, I think, to avoid self-flattery and self-loathing and thus to make genuine moral progress. But I think that we get something else

from self-scrutiny, even when that scrutiny does not yield particular self-knowledge. Specifically, in addition to providing us with an *awareness* of our moral agency, self-scrutiny can be a means of *affirming* our moral value.

To see what I have in mind, think about the value that we derive from intellectual inquiry. Sometimes, engaging in some intellectual pursuit yields knowledge. After years of carefully studying the *Groundwork for the Metaphysics of Morals*, for instance, we might finally come to understand what Kant means when he says in *Groundwork III* that he's reasoned in a circle. But sometimes our inquiry falls short; we may, after long and painful years of study, still find ourselves stumped. Nevertheless, there's still something morally valuable about intellectual inquiry that doesn't yield knowledge. Specifically, when we intellectually inquire into some topic, we affirm that that topic is worth trying to know. This commitment can be laudatory in itself, even if it doesn't bear fruit.

I think that unsuccessful self-inquiry has the same effect. We might never uncover the motives that led us to commit this or that action. But by looking closely at ourselves, we affirm that we're worth taking seriously. What we do *matters*—and so we need to pay attention, even if self-attention doesn't always yield self-knowledge.

Crucially, I don't think that this sense of self-worth is simply *revealed* to us through self-scrutiny. Rather, I think that the very act of self-scrutiny is a means by which we assert to ourselves that we're worthy of moral attention, in much the same way that the act of intellectual scrutiny affirms the value of its object. So, there's a funny thing that can happen when we set out

to know ourselves: even if we don't end up with particular self-knowledge, we may still end up with genuine self-worth.¹²

References

- Berg, Anastasia N. A. 2020. "Kant on Moral Self-Opacity," *European Journal of Philosophy* 28 (3): 567–85.
- Biss, Mavis. 2015. "Kantian Moral Striving," *Kantian Review* 20 (1): 1–23.
- Cholbi, Michael. 2022. *Grief: A Philosophical Guide*, Princeton, NJ: Princeton University Press.
- . forthcoming. "Respect, Self-Respect, and Self-Knowledge", *The Monist*.
- Cureton, Adam. 2017. "Kant on Cultivating a Good and Stable Will," in I. Fileva, ed., *Questions of Character*, Oxford: Oxford University Press, 65–77.
- Denis, Lara. 2001. *Moral Self-Regard: Duties to Oneself in Kant's Moral Theory*, New York: Routledge.
- . 2010. "Freedom, Primacy, and Perfect Duty to Oneself," in L. Denis, ed., *Kant's Metaphysics of Morals: A Critical Guide*, Cambridge: Cambridge University Press, 170–91.

12. This paper has had a long lifespan and has benefitted from significant feedback along the way. Thanks especially to Thomas E. Hill, Benjamin Bagley, G. Anthony Bruno, Gil Hersch, Daniel Hoek, Justin Horn, J. P. Messina, Ryan Preston-Roedder, Geoffrey Sayre-McCord, Oliver Sensen, and Philip Yaure. Thanks also to audiences at the UNC Workshop on Kant's Ethics and the North American Kant Society (Southern Division), and to an anonymous reviewer at this journal.

Fairbanks, Sandra Jane. 2000. *Kantian Moral Theory and the Destruction of the Self*, Boulder, CO: Westview Press.

Grenberg, Jeanine. 2005. *Kant and the Ethics of Humility*, Cambridge: Cambridge University Press.

———. 2018. “The Practical, Cognitive Import of Feeling: A Phenomenological Account,” in K. Sorensen and D. Williamson, eds., *Kant and the Faculty of Feeling*, Cambridge: Cambridge University Press, 41–66.

Halper, Aaron. 2023. “Kantian Self-Knowledge and the Opacity Thesis in Kant’s Doctrine of Virtue,” *Kantian Review* 28 (2): 185–200.

Kant, Immanuel. 1999. *Practical Philosophy*, ed. M. Gregor, Cambridge: Cambridge University Press.

———. 2001. *Lectures on Ethics*, eds. P. Heath and J. Schneewind, Cambridge: Cambridge University Press.

———. 2007. *Critique of Pure Reason*, trans. N.K. Smith, New York: Palgrave MacMillan.

———. 2012a. *Anthropology from a Pragmatic Point of View*, eds. R. Louden and G. Zöllner, Cambridge: Cambridge University Press.

———. 2012b. *Religion Within the Boundaries of Mere Reason*, eds. A. Wood and G. di Giovanni, Cambridge: Cambridge University Press.

MacKenzie, Jordan. 2018. “Knowing Yourself and Being Worth Knowing,” *Journal of the American Philosophical Association* 4 (2): 243–61.

Messina, James P. 2013. “Kant and the Priority of Self-Knowledge,” MA thesis, George State University, https://scholarworks.gsu.edu/philosophy_theses/135/.

- Moran, Kate. 2014. "Delusions of Virtue: Kant on Self-Conceit," *Kantian Review* 19 (3): 419–47.
- Moyer, Dean. 2006. "Unstable Autonomy: Conscience and Judgment in Kant's Moral Philosophy," *Journal of Moral Philosophy* 5 (3): 327–60.
- O'Hagan, Emer. 2009. "Moral Self-Knowledge in Kantian Ethics," *Ethical Theory and Moral Practice* 12 (5): 525–37.
- O'Neill, Onora. 1996. "Kant's Virtues," in R. Crisp, ed., *How Should One Live? Essays on the Virtues*, Oxford: Clarendon Press, 77–98.
- Papish, Laura. 2018. *Kant on Evil, Self-Deception, and Moral Reform*, Oxford: Oxford University Press.
- Roessler, Beate. 2014. "Autonomy, Self-Knowledge, and Oppression," in M. Oshana, ed., *Personal Autonomy and Social Oppression: Philosophical Perspectives*, New York: Routledge.
- Russell, Francey. 2020. "Kantian Self-Conceit and the Two Guises of Authority," *Canadian Journal of Philosophy* 50 (2): 268–83.
- Sticker, Martin. 2019. "Kant, Moral Overdemandingness and Self-Scrutiny," *Noûs* 55 (2): 1–24.
- Thomason, Krista. K. 2013. "Shame and Contempt in Kant's Moral Theory," *Kantian Review* 18 (2): 221–40.
- Ware, Owen. 2009. "The Duty of Self-Knowledge," *Philosophy and Phenomenological Research* 79 (3): 671–98.
- Wood, Allen W. 1999. *Kant's Ethical Thought*, Cambridge: Cambridge University Press.
- . 2009. "Duties to Oneself and Others," in T.E. Hill, ed., *The Blackwell Guide to Kant's Ethics*, West Sussex: Wiley-Blackwell, 229–51.