Kant on Moral Self-Opacity

Words: 4,984

Abstract
It has been widely accepted that Kant holds the ‘Opacity Thesis,’ the claim that we cannot know the ultimate grounds of our actions. The Opacity Thesis is, however, at odds with Kant’s claim that I act from the always potentially conscious representation of principles of action and that, in particular, in acting morally I act in consciousness of the moral law’s determination of my will. Thus, the Opacity Thesis threatens to render acting from duty impossible. To resolve the conflict, I, first, argue that a skeptical conclusion does not follow from Kant’s statements concerning the fallibility of our capacity for moral self-knowledge—the possibility of error or deception does not render knowledge in principle impossible. I then offer an interpretation of the kind of opacity that characterizes actions that are not performed from consciousness of the law.

I argue that self-opacity is a tendency intrinsic to the structure of our practical self-consciousness. It is precisely because the pursuit of our ends is self-conscious that a subject pursuing her own happiness must pursue the objects of her subjective inclinations and desires as if they were objectively good. Happiness, an indeterminate and self-undermining end consisting of the sum of all her subjective and contingent desires and inclinations, cannot be coherently pursued by an agent who is clear-sighted as to its real nature. Consequently, far from undermining the self-consciousness of practical reason, the self-opacity characteristic of bad action stems from the very standards internal to the will of a self-conscious agent. Thus I show how moral self-opacity in cases of morally bad action does not undermine the will’s self-consciousness but is born of it.
Introduction

The greatest obstacle to moral uprightness, for Kant, is neither figuring out what the right thing to do is (everyone, in the first place, knows what duty requires of them (KpV 5:36)), nor overcoming temptations to transgress (in knowing ourselves to be free we know that we could overcome any pathological temptation and do what is right (MS 6:380, R 6:47)). The greatest obstacle is rather self-deception, our peculiar ‘ability’ to hide from our own selves the grounds of our own actions. Repeatedly, Kant warns that when we have done the nominally right thing, we can be deceived about why it is we acted as we have—fancying ourselves to have been motivated morally, when in fact we were only pursuing our contingent desires. On the basis of Kant’s statements about the possibility of moral self-deception,¹ interpreters have widely attributed to him the “Opacity Thesis,” according to which we cannot know the genuine grounds of any given action—not only the actions of others, but even our own. Consequently, we cannot know whether any given action is performed merely in accordance with the moral law or from the genuine recognition of its authority.²

---

¹ Statements to this effect are found in G 4:407; R 6:51, 71, 87-88; MS 6:447; Anth 7:133, 143, 162. I turn to look at Kant’s treatment of moral self-opacity in detail in what follows.

² For example, Richard McCarty writes that, according to Kant, “we can never know whether this power [“the moving power of the moral ground of duty”] is the real determining cause of the choice to act” (Kant’s Theory of Action, 95); Allen Wood claims that “Kant thinks it is never easy (indeed, never even possible with certainty) to distinguish duty from self-love when we examine our inner motivation (Kant’s Ethical Thought, 30, see also 53, 201-202 and Kantian Ethics, 160); Onora O’Neill writes that “Kant insists that we do not know clearly what our maxims are on a given occasion. It is one thing to identify what the maxims of duty are—that the Categorical Imperative can achieve. It is quite another to ascertain what a particular agent’s maxim, even our own maxim, was on a particular occasion. Kant’s insistence on the opacity of the human mind and the limits of self-knowledge are extraordinarily strong” (“Kant’s Virtues,” 89); and Nancy Sherman claims that “what is opaque is not so much the ends or means we are pursuing, or accompanying effects—i.e., what we are doing—but why, or on what grounds” (“Wise Maxims/Wise Judging,” 50). Owen Ware offers the most recent lengthy treatment of the Opacity Thesis according to which: “one lacks direct cognitive access to the ground or purity of one’s disposition” and therefore “I cannot know, for example, whether my particular actions arise from conformity with the moral law or from some hidden self-interest” (“The Duty of Self-Knowledge,” 671-2, 674). Other
Interpreters have noticed that moral self-opacity presents potential obstacles to moral praxis. First, since moral uprightness is not a matter of doing the right thing but doing the right thing from recognition of the authority of the moral law, if we cannot know the grounds of our actions we cannot know the moral worth of our actions, and therefore can do little to address them. Accordingly, we run the risk of moral arrogance and

adherents of the thesis include Stephen Engstrom, “Conditioned Autonomy, 445-446; Beatrice Longuenesse, I, Me, Mine, 213; Claudia Bölö, Zurechnung bei Kant, 121; Falk Bornmüller, Selbstachtung, 172; Robert Louden, Kant's Impure Ethics, 78; Sven Bernecker, “Kant zur moralischen Selbsterkenntnis,” 163-181; and Patrick Frierson, Freedom and Anthropology in Kant's Moral Philosophy, 100. The earliest characterization of Kant’s alleged claim that moral self-knowledge is not possible as his “Opacity Thesis” that I am aware of is by David Cartwright, “Kant, Schopenhauer, and Nietzsche on the Morality of Pity,” 97. A recent dissenter is Laura Papish, who rejects skepticism about the possibility of moral self-knowledge. Papish draws attention to the fact that, in his theoretical philosophy, Kant insists that a limitation on knowledge (i.e., the impossibility of cognizing things in themselves) is compatible with an anti-skeptical epistemology; she adds that since self-deception involves consistent and careful diversion of attention from certain cognitions, self-deception presupposes self-knowledge of that which I am avoiding and to that extent self-knowledge is always “within her clear reach” (Kant on Evil, Self-Deception and Moral Reform, 165). While I share Papish’s ambition, namely, to provide an account that does justice both to the essential role of self-cognition as well as ubiquity of self-deception, my account differs decisively from hers. Papish, like Longuenesse and Ware, maintains that the propensity to self-deceive is a pathological or psychological one; by contrast I shall argue that it is a structural tendency born of the self-conscious character of our will.

Dina Emundts argues that the threat the Opacity Thesis makes to the possibility of moral evaluative judgment is sufficient to reject the Opacity Thesis: “In my eyes, it is essential for a philosophy of morality that it also give us a criterion for evaluating actions that have been committed. Thus, Kant has to claim that self-knowledge in the sense of knowledge about our own motives is possible” (Emundts, “Kant’s Ideal of Self-Knowledge, 92).

Considerations of this sort lead Sebastian Raedler and Sven Bernecker to reject the claim that we may not know our particular maxims because this would render it impossible to evaluate them for permissibility, which, on their accounts, is necessary for acting morally. See Raedler, Kant and the Interests of Reason, 128-133; Bernecker, “Kant zur moralischen Selbsterkenntnis,” 164. However, the focus of Kant’s concern in his remarks on moral self-opacity is not the difficulty of knowing a particular maxim but the possibility of being mistaken about whether or not we were genuinely motivated morally,
pernicious moral self-certainty. Interpreters have also worried that the thesis appears to conflict with Kant’s claim that one has a duty to know oneself, as well as with his accounts of moral practices that depend on the possibility of evaluating others’ actions as good (like moral education or so-called appraisal respect). For all of that, the thesis, which seems to encapsulate an important truth about our moral predicament, has been favorably viewed as a sign of Kant’s moral-psychological sophistication.

The Opacity Thesis is not, however, only an obstacle to those tasks in which individuals must know the moral worth of their actions or those of others. It moreover undermines the very possibility of acting from duty, an exercise of our will, which, according to Kant, requires knowledge of the moral law as the determining ground of the will. For Kant, the consciousness of one’s principles of action is constitutive of the exercise of our power to act. In exercising my will I act from the (always potentially conscious) representation of principles; in the case of acting morally I act from consciousness of the moral law’s determination of my will. Knowing the moral law to be the ground of one’s willing is thus a condition on acting morally. If the opacity thesis is right and one cannot know the ground of one’s will, then one cannot act from duty. In this paper I propose an account of Kantian moral self-opacity that preserves the possibility of moral self-knowledge and thus morally good action as well as illuminates the way in which moral self-deception and self-opacity are not merely motivated by contingent, pathological moral vanity but are structurally intrinsic to the will of a self-conscious agent.

1. Moral Self-Opacity

---

i.e., knowledge of the ultimate ground of our actions. For an argument to the effect that it is not specific intentions that we cannot know but our fundamental maxim (or Gesinnung) see Henry Allison, Kant’s Theory of Freedom, 93. It is the opacity of the ultimate ground of our will that, I shall argue, poses a profound challenge to the possibility of acting from duty. Raedler and Bernecker both accept that, according to Kant, we cannot access the ultimate grounds of our actions, and therefore they accept the aspect of the Opacity Thesis that I mean to dispute in this essay.

5 See Ware, “The Duty of Self Knowledge”; Wood, Kant’s Ethical Thought, 201.

6 See Hakim, “Kant on Moral Illusion and Appraisal of Others.”
It is first important to distinguish Kant’s account of moral self-opacity from his claim that we cannot know the moral worth of other people’s actions.\(^7\) Kant gives substantively different grounds for both. The difference turns on the radically different perspectives we can have on action: the perspective we can take on the actions of others, when we consider those as facts of experience, and the perspective we take on our own actions as exercises of our rational wills. In particular, Kant’s characterization of moral self-opacity is not an instance of the general epistemic restriction against having empirical knowledge of non-empirical phenomena.\(^8\) Rather, Kant’s account of moral self-opacity concerns a difficulty in knowing the grounds of our own actions from a perspective from which such knowledge is still in principle possible. Moral self-opacity concerns the fallibility of a non-empirical, properly moral variety of knowledge. As such, it is a specific form of moral failure.\(^9\)

Consider the following statement from the *Groundwork*:

It is indeed sometimes the case that with the keenest self-examination we find nothing


\(^8\) Pace interpreters like Allen Wood who claims: “Kant is thus very much in agreement with Nietzsche’s critique of “naïve empiricism”: the “inner” world of our sensations and feelings is even less trustworthy and more “phenomenal” than the world of external objects” (Wood, *Kant’s Ethical Thought*, 201-202). Wood goes on to argue that for the purposes of understanding Kant’s account of self-opacity, there is no need to bother with his “metaphysical theory of freedom, which locates free agency in the intelligible world.” (*Ibid*.). Stephen Engstrom offers a different explanation for moral self-opacity that nevertheless grounds moral self-opacity in the constitutive limitations of our empirical knowledge. Engstrom claims that we cannot know the ground of our action because such knowledge would require knowing what I would have done in every possible circumstance and such knowledge is not available in experience: “[t]he impossibility that Kant refers to…rests on the incontestable fact that none of us are in position to know with certainty how we would act in every possible circumstance” (Engstrom, “Conditioned Autonomy,” 446).

\(^9\) Most interpreters recognize self-opacity to be a special case of distinctively moral blindness. Ware, for example, identifies it as a special kind of opacity whereby “I can never be certain of the moral purity of my actions because of my deeply selfish nature” (Ware, “The Duty of Self-Knowledge,” 674); Longuenesse calls it a “motivated blindness to the grounds of one’s actions” (*Longunesse, I, Me Mine*, 205).
besides the moral ground of duty that could have been powerful enough to move us to this or that good action and to so great a sacrifice; but from this it cannot be inferred with certainty that no covert impulse of self-love [geheimer Antrieb der Selbstliebe], under the mere pretense of that idea, was not actually the real determining cause of the will; for we like to flatter ourselves by falsely attributing to ourselves a nobler motive, whereas in fact we can never, even by the most strenuous self-examination, get entirely behind our covert incentives [die geheimen Triebfedern]. (G, 4:407, emphasis mine)

Here Kant is not denying that we can have any access to the grounds of our actions; on the contrary, he admits the possibility of discovering some motives, but not others. The trouble is that the agent cannot see into the “depths” of her heart to “be quite certain” or “get entirely behind” her non-moral incentives. In the “self-examination” Kant has in mind the subject is not simply restricted to the observation of natural changes in the world, but can find “the moral ground of duty.” The case is therefore one in which not only does the agent know that the action is in accordance with the moral law, she also knows that it was her duty to perform it. Only thereby could the agent be in a position to “find” the “moral ground of duty” by means of self-examination and deceive herself that the moral ground is what in fact motivated her.

This and similar passages\(^{10}\) suggest that in characterizing moral self-opacity, Kant is describing a perspective from which it from which one’s motives are not inaccessible but from which they may be hidden or obscured.

Statements like these have led interpreters to attribute the Opacity Thesis to Kant, the claim that it is in principle impossible to know the grounds of one’s own actions.\(^{11}\) As we shall see shortly, however, the Opacity Thesis cannot be reconciled with Kant’s account of the self-consciousness that is constitutive of the exercise of our rational capacities in general and our will

\(^{10}\) See especially the following from the Metaphysics of Morals: “a human being cannot see into the depths of his own heart so as to be quite certain, in even a single action, of the purity of his moral intention and the sincerity of his disposition, even when he has no doubt about the legality of the action”. (MS, 6:392). The agent has knowledge of his duty (for otherwise he could not assess the legality of the action) and is aware of a moral intention, but cannot be quite certain whether this moral intention is “pure”, i.e., whether there existed another motive besides. See also, (R, 6:38).

\(^{11}\) See footnote 3 for references.
in particular. It thereby threatens the core of Kant’s account of morally worthy action—the possibility of knowledge of the moral law as the ground of one’s will.

3. Self-Opacity and Self-Consciousness

The Opacity Thesis renders illusory our claim to know the grounds of our own actions. In what follows we shall see that it thereby threatens the very possibility of morally worthy action. This is because for Kant, rational representational capacities are essentially self-conscious. The self-consciousness of the will, in particular, implies that to act on a principle is to act in consciousness of acting on that principle. Our rational faculty of desire is a power to act by means of representations, but it is not merely that: it is the capacity to act from representations of rational principles (cf. G 4:412, 428, 459). Since it is constitutive of our representative capacities that we can always bring our representations explicitly to mind (B131-132), and since we do not act merely in accordance with laws but from the representation of them, we can bring to consciousness the principles upon which we act. In turn, this general account underlies the well-known Kantian thesis that insofar as acting well requires acting from consciousness of the principle which is the moral law, one must not only do as one should, but one must do as one should because one knows one ought to so act—and therefore through knowledge of the moral law as the principle of one’s action. In acting well, i.e., from the consciousness of the moral law, the agent therefore necessarily knows the moral law as the ground of her action.

Kant explicitly claims that morally worthy action requires knowledge of the grounds of one’s action. Consider the following claim from the *Groundwork*:

What I cognize immediately as a law for me I cognize with respect [Was ich unmittelbar als Gesetz für mich erkenne, erkenne ich mit Achtung], which signifies merely consciousness of the subordination of my will to a law without the mediation of other influences on my sense. (G 4:401fn)

Kant states that to act from the law is to act in consciousness of the subordination of the will to the moral law, i.e., in consciousness of the moral law, unmediated, as determining one’s will.
This consciousness, as his use of “cognize” [erkenne] is a form of knowledge.\textsuperscript{12} To recognize the authority of the law is to know it as the law, i.e., as a principle that has the objective authority to command me to act. This is not a form of theoretical cognition, but a form of practical knowledge. Morally worthy action requires not only cognition (knowledge) of the moral law as determining the will, or what is the same, as the ultimate the ground of one’s action but also consciousness that the subordination of the law to the will occurs without the mediation of other sensible influences. To act from the moral law, requires consciousness of acting from it. Likewise Kant writes in the second Critique, “The moral disposition is necessarily connected with consciousness of the determination of the will directly by the law” (\textit{KpV} 5:116) and adds shortly after that morally worthy action requires “consciousness of direct necessitation of the will by the law” (\textit{KpV} 5:117).

The Opacity Thesis, we have seen, renders such cognition illusory—for it alleges we cannot know what it is that ultimately determined my will. Thus, the Opacity Thesis not only stands in conflict with the central Kantian principle that any representation can in principle be brought to consciousness; it threatens to undermine the \textit{very possibility of performing morally worthy actions}.

\section*{4. Moral Self-Knowledge Disjunctivism}

The first step in overcoming this interpretative and philosophical impasse is to reject the idea that Kant’s account of moral self-opacity implies the Opacity Thesis.

To this end, it would be useful to appreciate the way in which the structure of the problem of moral self-opacity as it has emerged resembles one familiar from contemporary debates in epistemology (and, in particular, the philosophy of perception) about the possibility of knowledge in the face of the potential for error. The general structure of the problem is as follows. It seems that the undeniable fact of error—exemplified in cases where a subject takes

\textsuperscript{12} The topic of what practical knowledge consists in for Kant is of course a vast one. (See especially Engstrom, \textit{The Form of Practical Knowledge}, for a thorough and illuminating meditation on the topic). For my purposes in this essay, I restrict myself to the claim that it is precisely the sort of knowledge that is denied by the Opacity Thesis, the knowledge of the grounds of our own will, that Kant affirms is intrinsic to morally worthy action.
herself to know something on the basis of some evidence but is mistaken—apparently entails that the relevant capacity to know is essentially fallible. This is because, the argument goes, if error is possible, its possibility cannot be excluded on any given occasion, but precisely such exclusion is necessary for knowledge. Knowledge implies the exclusion of the possibility of error and so if error cannot be excluded a genuine claim to knowledge is never warranted. Therefore, a successful exercise of the capacity to know is, in principle, impossible.

I suspect that a logic similar to the one I outlined here has guided interpretations of Kant’s pronouncements regarding moral self-opacity. Applying the same sceptical reasoning to the case of knowledge of the grounds of one’s actions, the Opacity Thesis presupposes that the fact that error is possible, i.e., that I may be falsely convinced my motives are pure, justifies the conclusion that moral self-knowledge is in principle impossible.\textsuperscript{13}

However, the mere possibility of error, i.e., fallibility, does not render the capacity for knowledge in general, and the capacity to know the grounds of one’s own actions in particular, essentially inadequate to their task. To avoid the sceptical conclusion, I suggest we consider Kant’s account of moral self-knowledge and self-opacity along so-called “disjunctivist” lines.\textsuperscript{14}

In the philosophy of perception disjunctivist accounts turn on the idea that we can avoid the sceptical conclusion, if we reject the “highest common factor” (HCF) account of experience—the idea that the content and epistemic significance of experience can be identical in veridical and non-veridical cases. To reject the HCF account is to recognize that the subjective inability of the perceiver to tell the ‘bad cases’, e.g., of hallucination, visual error and so forth, from the ‘good ones’, \textit{from within the perspective of the bad cases}, does not imply any objective identity between the cases. Once the disjunctivist is granted that good and bad cases (cases of

\textsuperscript{13} Importantly, while Kant’s statements leave little doubt that he thought moral self-opacity is a serious and ubiquitous threat, he never claims that it is not possible to know the grounds of our actions. In other words, Kant is firmly committed to the fallibility of our moral self-knowledge but he never affirms the sceptical conclusion that is the Opacity Thesis.

\textsuperscript{14} Disjunctivist accounts of knowledge trace to J. M. Hinton, and today are closely associated with John McDowell. See Hinton, “Visual Experiences,” 217-227; and \textit{Experiences}, esp. Part II; McDowell, “Criteria, defeasibility, and knowledge,” 455-79. For a useful overview see Alex Byrne and Heather Logue’s “Introduction” to their collected volume, Byrne and Logue, eds., \textit{Disjunctivism}, vii-xxix.
knowledge and error) need have no factor in common over and above it seeming to the subject, in both cases, that things are a certain way, she has what she needs to deny the skeptical conclusion. Acknowledging that a capacity to know is fallible does not entail that it is not in fact a capacity for knowledge at all. In particular, it does not follow that there are no good cases in which the subject making a judgment on the basis of evidence is warranted in their knowledge claim. For example, in a case of genuine perception I have sufficient warrant for my knowledge claim.

Similarly, I propose, the fact that an agent may be deceived about the purity of her moral motivation does not imply that knowledge of the grounds of one’s action is impossible. While it may seem to the subject that she is motivated morally both in cases when she is and when she is not, she is justified in her knowledge claim in some of those cases and not in others. A claim to knowledge does not require evidence for the exclusion of error, or proof that one is not in fact in the bad case. In cases of acting from duty the agent is just thereby justified in her claim to know the ground of her action. Thus, fallibility does not imply that moral self-opacity is an essential feature of any and all exercises of our will. The disjunctivist argument thus secures the possibility of the knowledge of one’s grounds in the case of action from duty.

We find therefore room both for the possibility and ubiquity of moral self-opacity as well as for the kind of moral self-knowledge that is, as we have seen, required in cases of morally good action, cases in which the subject must act in consciousness of the moral law as the ground of their will. However, the introduction of the theme of the self-consciousness of the representational capacity puts pressure on—and thus is an occasion for illumination of—the very possibility of self-deception.

5. Self-Deception and Self-Consciousness

Recognizing the self-consciousness of our representational capacities raises a further question that is not answered by the disjunctivist argument offered in the previous section: how should we understand the possibility of self-opacity in cases of action that are not performed from duty?

Even if we are persuaded that moral self-opacity is not a general feature of all exercises of the will but is only characteristic of bad action, my lack of awareness of the deception poses a challenge the perceptual case does not. How can I be blind to the grounds of my own action? It is
after all, *my* subjective motivations that I am gratifying by means of the bad action, the *self*-love that is driving me is the love of *my own* self. How are they mine if I cannot recognize them as my own?

It is tempting to think of self-deception as an intentional action that an agent performs. This in turn makes attractive readings that explain self-deception by reference to a material desire, such as moral vanity, which motivate the intentional action in question: that of hiding the truth about one’s moral motivations from oneself. The paradigmatic case of self-deception is one whereby the subject is concerned enough with the demands of the moral law that it would be sensibly painful for her to be forced to acknowledge that although she acted well, she did so selfishly. The subject, in other words, has a pathological (non-moral, extra-rational) preference to consider herself as acting from the moral law or is morally vain. While the account captures well the familiar desire of wanting to appear better than we in fact are, it does not do much to illuminate the possibility of presenting a false appearance of our moral character to ourselves. In fact, it renders it incoherent. To see this, suppose I have a sensible motivation to hide the grounds of my action from myself: how am I not aware of the end and principle that I act on to hide the ground of my action from myself? It appears we must now posit an additional sensible motivation to hide the secondary sensible motivation that motivated me to hide the original ground of my action, and another to hide that one, and so on.

Underlying this threat of regress is an over extension of Kant’s concept of a “sensible motive.” On this interpretation a sensible motive can motivate ordinary action as well as motivate the ‘act’ of hiding from oneself the ground of one’s action. But is the latter ‘act,’ the so-called act of self-deception, an act in the same way as an ordinary action is? Is it, in particular, the kind of thing that we can be ‘sensibly motivated’ to do? This does not seem right.

15 Beatrice Longuenesse expounds a reading along these lines: “just as we have sensible motives to act against the moral law, we also have sensible motives (motives of self-love, escalating into self-conceit) to deceive ourselves about the motivation of our actions” (Longuenesse, *I, Me, Mine*, 214.) A similar interpretation is implicit in Jeanine Grenberg: “Self-deception becomes a tool whereby we construct an image of ourselves more pleasing to ourselves” (Grenberg, “Self-Deception and Self-Knowledge,” 164). Likewise, Papish claims that self-deception involves “rationalization” which “is what occurs when we introduce a desirable cognition or hoped for justification into the reasoning process” (Kant on Evil, Self-Deception, and Moral Reform, 74). See also Ware, “The Duty of Self-Knowledge,” 674.
For Kant, to be ‘sensibly motivated’ is to act in the expectation of pleasure in the reality of an object in accordance with a practical principle (KpV 5:21). A sensible motive is a way in which our faculty of desire, our faculty to act, can be determined, namely, in the expectation of gratification in the attainment of an object. But the determination of what the moral agent is or is not aware of in her determination to action is not itself another “action.” Sensible motives do not determine what is available to consciousness, they influence the will into pursuing particular ends. The difference is radical. While we can determine ourselves to pursue a particular end in the expectation of gratification in its realization, we cannot consciously determine ourselves to remove, from consciousness, other sensible motives in order to gratify a pathological interest in our own moral superiority.

6. Structurally motivated self-deception

I suggest that self-opacity is not sensibly motivated but is rather a tendency intrinsic to the structure of our practical self-consciousness. Self-deception and thus self-opacity is called for as a response to the essential difficulty of bringing to consciousness the ultimate motivating grounds of an action performed merely for the sake of gratification in a particular object. This difficulty is not due to the fact that, in acting for the sake of self-love, we are dealing with extra-rational feelings that are thereby, for some reason, hard to represent. As we have seen, even in such cases there exists a representation of both an object aimed at and a principle by means of which one intends to secure it. Rather, acting for the sake of self-love resists conscious representation because there is an essential indeterminacy in the concept of the ground to which all such motivations belong, namely, what Kant calls “happiness.” Because happiness is an indeterminate and worse, always potentially self-contradictory end, in pursuing a particular contingent end for the sake of happiness (and not merely in the pursuit of a moral one) I am thereby committed to an incoherent project. An incoherent project is one that the agent cannot represent herself as coherently pursuing. It is this impossibility of representing happiness as an end, I will argue, that both enables and promotes self-opacity.

According to Kant, there are three ways in which the capacity of desire can act in accordance with the representation of laws or principles (G 4:412): imperatives of skill, prudence and moral. When an agent pursues a particular end not for the sake of the moral law, one is guided by
imperatives of prudence (G 4:414). These assertoric hypothetical imperatives represent “the practical necessity of an action as a means to the promotion of happiness” (G 4:415). Hence, they are dependent for their intelligibility on the concept of happiness. Happiness indicates neither a determinate object nor a distinct form of action. It denotes “the sum of satisfaction of all inclinations” (G 4:399) and this means that it cannot be determined concretely: “one can form no determinate and sure concept of the sum of satisfaction of all inclinations under the name of happiness” (G 4:399). The concept of happiness cannot can serve as a unifying end of action, and one cannot therefore conceive of a coherent principle that aims at its achievement.

Indeed, it is, Kant writes, “such an indeterminate concept that, although every human being wishes to attain this, he can still never say determinately and consistently with himself what he really wishes and wills” (G 4:418, emphasis mine). The difficulty of determining how to pursue happiness is multifaceted. First, securing happiness would require omniscience: since happiness depends on the consequences of our actions, we would need to be able to predict their effects. Since we possess no such power, in pursuing one goal we cannot secure against the possibility that it may bring about undesirable consequences in its wake:

[F]or the idea of happiness there is required an absolute whole, a maximum of well-being in my present condition and in every future condition. Now, it is impossible for the most insightful and at the same time most powerful but still finite being to frame for himself a determinate concept of what he really wills here. (G 4:418).

Kant goes on to list various unforeseen consequences of pursuing presumptively happiness-promoting ends like riches, a long life, health, etc. And concludes, “In short, he is not capable of any principle by which to determine with complete certainty what would make him truly happy, because for this omniscience would be required” (G 4:418). The first difficulty is therefore the fact that that which “brings true lasting advantage” is, as Kant recognizes, “always veiled in impenetrable obscurity” (KpV 5:26). Secondly, it is not only that the satisfaction of one’s current desires might lead one astray in the future, but that since one’s inclinations and subjective desires are only known empirically, one can always find oneself desiring new, different, potentially contradicting things: “it cannot be cognized a priori of any representation of an object, whatever
it may be, whether it will be connected with *pleasure* or *displeasure* or be *indifferent*” (*KpV* 5:21).

Moreover, even presently known desires may conflict with one another. Any agreement between one’s contingent desires would always be merely accidental and contingent.\(^{16}\)

Consequently, Kant goes as far as to strip imperatives of prudence from their ‘imperative’ title. They cannot *command with necessity* since they cannot present actions “objectively as practically necessary”, (*G* 4:418), which imperatives of skill do relative to any chosen end, and categorical imperatives do absolutely.

The end of acting from the motive of “happiness,” or the sum of all my subjective, contingent desires and inclinations, is an indeterminate end, at any moment potentially self-contradictory, and therefore impossible to systematically pursue.\(^{17}\) Such an indeterminate conception of our ultimate end does not offer standards for its attainment and therefore cannot guide practical reasoning. Therefore, since no finite rational being is capable of a principle by which to determine what would make her happy, no such being, when her will is determined by considerations of self-love can in fact be determined to action by a principle which articulates what is necessary to attain her end.

This, I propose, is the root of self-deception. It is precisely because the pursuit of her ends is self-conscious that the agent’s pursuit of happiness must treat the objects of her subjective inclinations and desires *as if* they were not aimed at the indeterminate end that happiness in fact is, *subjectively* desired, but as though they were *objectively* good. The agent does not, and cannot, view herself as undertaking the impossible task of harmonizing her contingent desires but pursues her ends *as if* they were non-accidentally unified in a single whole, as if she is guided by an ideal of objective practical necessity, not subjective contingency.

The failure involved in moral self-opacity is thus not a matter of blindness to the presence of a particular desire, end, or an articulated practical principle upon which one is acting, and

\(^{16}\) Cf. *G* 4:418-419; *KpV* 5:36-37.

\(^{17}\) This is not to deny that happiness is desired and someone may try to achieve it by devising principles in order to maximize the satisfaction of their desires; rather, it is to suggest that there is no principle available which can articulate a necessary relation between a course of action and the fulfillment of this end.
which is impossible to ever bring to consciousness. It is rather a matter of misrepresenting, in the pursuit of a particular end, what it is that renders the end attractive to oneself. The end, which is dependent on a contingent desire, is misrepresented as the concrete realization of one’s moral duty.

Thus, moral self-opacity is not in the first place an expression of contingent pathological moral vanity. It is rather a tendency internal to the operation of a rational faculty of desire: the tendency to pursue the agent’s happiness as if it were the good.

This interpretation accords perfectly with Kant’s explicit characterization of moral self-opacity, as a condition whereby a “covert impulse of self-love,” serves as the real determining cause of the will “under the mere pretense” of the “idea” of “the moral ground of duty” (G 4:407, emphasis mine). The impulse of self-love is not simply hard to detect, it is disguised as something other than itself, namely, as moral motivation. One pursues a particular end under the pretense of the moral law, i.e., as if one is acting from the ground of duty.\textsuperscript{18}

This interpretation derives additional support from its agreement with Kant’s difficult claim that even when acting morally badly agents cannot consciously reject the moral law. According to Kant, the agent cannot ever act in consciousness of their rejection of the authority of the moral law. So that while agents certainly do evil (indeed we have seen Kant insist that it is not certain they have ever done anything else), the nature of their transgression, i.e., their disregard for the law, cannot be brought to consciousness. Kant rejects the conscious choice of evil, which he calls diabolical or devilish [\textit{ein teuflischer Wesen}] (R 6:35), as an incoherent form of moral failure. Thus, he writes in the \textit{Religion}:

To think oneself as a freely acting being, and yet as released from the law commensurate with such a being (the moral law), would be tantamount to thinking a cause operating

\textsuperscript{18}Laura Papish recently proposed a general account of Kantian self-deception. According to Papish, deception involves selective attention to evidence by way of a deflection of attention from an original cognition through a process of “rationalization.” (Papish, \textit{Kant on Evil, Self-Deception, and Moral Reform}, 73). At this level of generality my account is in agreement with hers. Nevertheless, Papish sees rationalization as motivated purely pathologically (cf. fn16), an idea I wish to reject.
without any laws (for, on account of freedom, determination according to natural laws drops out); and this is contradictory. (R 6:35.)

Because rejecting the law would involve thinking oneself both as free and released from the law, Kant thinks it is simply impossible. This rejection would undermine the agent’s very agency. The problem with alleging that an agent can choose to deliberately subvert the moral law lies in the impossibility, from the standpoint of the practical agent, of thinking oneself at once as a freely acting being—a condition for exercising choice—and at the same time as released from the law. A certain level of moral self-deception is thus required for the agent to reject the authority of the law at all.

Thus we see that the kind of moral pretense that is involved in self-deception does not answer a merely pathological predilection—the state it describes is not merely sensibly desired or hoped for. Moral self-opacity is structurally motivated by the intrinsic self-consciousness of our will and the difficulty of practically representing oneself as acting for the indeterminate end of happiness—for to act is to act for an end, and an indeterminate end is ultimately no end at all.

7. Conclusion

---

19 Cf. “The human being (even the basest), no matter in what maxims, does not, as it were, in a rebellious manner renounce the moral law (by revoking his obedience to it)” (R 6:36).
20 Notice especially that what renders the choice of evil unintelligible is not whether it is chosen as a means or as an end but whether it is chosen knowingly.
21 Kant’s denial of the possibility of diabolical evil has come under criticism from John Silber, “The Ethical Significance of Kant’s Religion,” cxxix, and “Kant at Auschwitz.” This criticism has been reiterated by R. J. Bernstein, Radical Evil: and most recently Claudia Card, The Atrocity Paradigm. Their objection is empirical: they insist that in fact people have exhibited and exercised the capacity to deliberately reject the moral law. Putting aside the question of the validity of such empirical claims, the very idea that one can disprove Kant’s claim with empirical evidence misses the thrust of Kant’s argument. Kant’s claim that agents cannot disobey the law deliberately is conceptually grounded in his account of moral personality. On this see Wood, Kant’s Moral Religion, 212-215 and “Kant and the Intelligibility of Evil,” 152-155; and Allison, “Reflections on the Banality of (Radical) Evil,” 149.
22 For a thorough overview of the relation between evil and self-deception in Kant see Papish, Kant on Evil, Self-Deception, and Moral Reform, esp. Chapters 3 and 6.
The reconstruction I have presented shows how we can unify disparate elements in Kant’s text: the ubiquity of moral self-deception and self-opacity, Kant’s general commitment to the self-consciousness of rational, representational faculties; his particular account of the way in which practical knowledge of the ground of one’s action is intrinsic to moral action, and his account of happiness. I hope that at the same it provides a general indication of how philosophical accounts of mind and action, which are grounded in the self-consciousness of our rational capacities, can not only make room for cases of failures of self-consciousness, but can themselves provide the resources to deepen our understanding thereof.

The interpretation also helps us make sense of another curious and apparently self-contradictory feature of Kant’s overall attitude to moral self-opacity and self-knowledge, namely, Kant’s warning against excessive moral self-scrutiny. While Kant is eager to warn his readers that moral self-deception and therefore moral self-opacity are a genuine threat to moral practice, Kant likewise cautions against the attempt to overcome the possibility of opacity by means of self-examination. “All self-scrutinizers fall into the gloomiest hypochondria” (VA 25:863). The “descent into the hell of self-cognition” (MS 6:441), as Kant vividly calls it, can do as much to jeopardize one’s moral constitution as to fortify it. The interpretation of moral self-opacity that I propose helps us see why Kant might have been right to warn against attempts to morally scrutinize oneself along these lines: it is pointless. Nothing in Kant’s account suggests that moral self-deception can be overcome by means of our voluntary efforts and the interpretation I propose makes clear why: to go searching after covert motives is to misunderstand the very sense in which they are covert. Moral self-opacity is not another amoral act to which we are tempted by pathological incentives. It is a structural self-blinding motivated by our intrinsic tendency to strive to be adequate to the standards of our own will. What is left for us to do is not to go hunting after secret motives but rather to practice moral humility: to suspect our own claims to moral purity, in our actions and our judgments of others.
Bibliography and Abbreviations

Kant’s works

In the case of the *Critique of Pure Reason*, I follow the standard practice of referring to the 1781 (A) and 1787 (B) editions. For all other texts, citations appear in the order of abbreviation, volume number, and page number from the *Akademie Ausgabe* (AA), *Kants Gesammelte Schriften*, edited by Königlich Preussische Akademie der Wissenschaften (29 vols. Berlin: de Gruyter, 1900–). Unless otherwise stated, all translations come from *The Cambridge Edition of the Works of Immanuel Kant*, edited by Paul Guyer and Allen Wood (Cambridge: Cambridge University Press, 1992–).

*Anth.* Anthropologie in pragmatischer Hinsicht (AA 7), *Anthropology from a Pragmatic Point of View*. Translated by Robert B. Louden.

*G* Grundlegung zur Metaphysik der Sitten (AA 4), *Groundwork for the Metaphysics of Morals*. Translated by Mary Gregor.

*KpV* Kritik der praktischen Vernunft (AA 5), *Critique of Practical Reason*. Translated by Mary Gregor.

*MS* Die Metaphysik der Sitten (AA 6), *The Metaphysics of Morals*. Translated by Mary Gregor.


Other sources


Longuenesse, Béatrice. *I, Me, Mine: Back to Kant, and Back Again*. Oxford University Press, 2017. [*I, Me, Mine*]

Louden, Robert B. *Kant’s Impure Ethics: From Rational Beings to Human Beings*. Oxford University Press, 2002. [*Kant’s Impure Ethics*]


