

Kant on Reflection and Virtue

Response to Colin McLear, Francey Russell, and David Sussman

Melissa Merritt

I want to thank the Society for German Idealism and Romanticism, and especially Keren Gorodeisky, for organising this series of reviews. I would also like to thank Colin McLear, Francey Russell, and David Sussman for taking the time to read my book and respond with such great care. Your comments were extremely thoughtful, and pushed me to think harder about some of my claims in the book, and also about where I want to go from here.

1.

One of the puzzles that drove my research early on was the enormous complexity in the textual record on reflection. And yet the secondary literature on reflection was somewhat over-focused, I thought, on a couple of paragraphs in the *Jäsche Logic*, which was not a text that Kant wrote or prepared for publication himself. The *Jäsche Logic* passage was about three mental acts in the formation of concepts, one of which is called reflection. Before long, commentators were taking the *judgment of reflection* at issue in the third *Critique* simply to be the reflection mentioned in the *Jäsche* paragraph on concept formation.¹ This seemed like a mistake to me (but not one that I felt able to correct in my book). Another important text on reflection is found in the first *Critique*, in the Amphiboly chapter. It discusses reflection in ways that I initially connected with Kant's conception of enlightenment, and eventually came to see as part of the "humanist" side of Kant's logic² — specifically the "applied," rather than pure, general logic that deals with the messiness of the human epistemic condition, such as our susceptibility to prejudice, as well as with other topics that interested me, like the direction of our attention in experience. Kant develops these concerns in his later writings, chiefly in the

¹ Henry Allison, *Kant's Theory of Taste* (Cambridge: Cambridge University Press, 2001), is representative of this tendency.

² As Huaping Lu-Adler observes in *Kant and the Science of Logic* (New York: Oxford University Press, 2018), Kant is taking up a humanist tradition as he develops applied logic.

Anthropology. While I wanted to be able to find my way, on principled systematic terms, through the complex textual record on reflection, ultimately I realised that my chief interest was to understand what Kant's talk about a requirement to reflect means for us at the ground level.

To sort out the complexity of the textual record, I drew a distinction: namely, between senses of reflection that seem to concern what is constitutively involved in thinking as such, on the one hand, from a normative sense of reflection that Kant characterises, variously, as a requirement or even a duty. Colin McLear is not convinced by this fundamental distinction that I propose. He also notes that the record is even more complex than I acknowledge, since there is a passage in the First Introduction to the third *Critique*, where Kant claims that reflection “goes on even in animals, though only instinctively” (20:211). I set that passage entirely aside: I don't know how Kant can talk with such confidence about the mental lives of non-rational animals — and, at any rate, I am concerned with reflection as an expression of rational mindedness.

Some of the earlier literature on reflection identified reflection with *comparison*, based partly on remarks from the *Jäsche* passage on concept formation.³ I believe that McLear endorses this identification when he points to places where Kant seems to explain reflection as a kind of comparison: e.g., in transcendental reflection, one compares a given representation with the cognitive capacity from which it should arise; and in logical reflection, one compares representations for the purpose of generating a concept. However, I think that comparison is simply too indeterminate a notion to give us a handle on what reflection is. As I argue in the book, if one invokes the notion of comparison, one must also invoke — at the very least — a standard of comparison: x and y are compared in length, or wisdom, or hue, and so on. And then we have to consider the point of making the comparison, for which we must invoke some *end* or governing principle. So, I just do not think it will do, in the face of such a complex textual record, to simply say that reflection is comparison. Comparison is involved in reflection, as I said in the book; but they are not the same thing.

McLear discusses Kant's notion of transcendental reflection from the Amphiboly chapter. I set that aside in the book, although I did discuss it in a 2015 paper (part of which appears in the book).⁴ So I am not going to take up that aspect of his comments at much length.

³ See, e.g., Béatrice Longuenesse, *Kant and the Capacity to Judge* (Princeton, NJ: Princeton University Press, 1998).

⁴ Melissa Merritt, “Varieties of Reflection in Kant's Logic,” *British Journal for the History of Philosophy* 23, no. 3 (2015): 478-501. I am not going to try to defend my reading of *transcendental* reflection from that paper

However, let me briefly say that one of my goals in that paper and the book is to direct attention to the overlooked aspect of Kant's logic: namely his conception of *applied*, rather than pure, general logic. As I mentioned at the outset, Kant directly nods to applied logic, and its dealings with our susceptibility to prejudice, at the start of the Amphiboly. That is what is on his mind when he talks about transcendental reflection. According to applied logic, prejudices are fundamentally expressions of a kind of *epistemic self-conceit*. The point comes out nicely in the Blomberg logic (one of the earlier records of Kant's logic lectures, where his interest in the logical humanist tradition is developed at particular length): "a prejudice is indeed nothing other than the mere desire to want to judge, but without the proper acuity or reflection" (24:187).⁵ Now I do think that McLear may be overlooking the significance of this sort of point from applied logic; so I will try to sketch where that occurs and what its implications are.

Kant with great consistency names three sources of prejudice: *custom*, *inclination*, and *imitation*.⁶ McLear is right to point out that they "constitute principles for associating representations." In this context, McLear suggests that an ambiguity in Kant's notion of *judgment* makes mischief when it comes to Kant's applied-logic claim that "all judgments require reflection" (A261/B317). McLear suggests that Kant sometimes takes *judgment* "to denote the act of contentful, truth-functional representation (what we might now describe as the grasp of a propositional content)" and yet "at other times uses it to denote the *acceptance* of the truth of such a content." I haven't worked through the textual record with this matter in mind, but for the sake of argument I will accept this observation. Now, McLear suggests that the sense in which one "judges" without reflection — *as when Kant explains prejudice as a judgment without reflection*⁷ — is at most the first sense just distinguished, i.e. grasping a truth-functional content. The sense of judgment that involves endorsement or acceptance of that content can occur only "in virtue of epistemic and logical laws and not via associative mechanisms tied to sensibility, such as custom, inclination, or imitation." Thus McLear says that the prejudice is not really judgment, because it occurs in us owing to the force of these associative mechanisms.

here: first, I am not entirely sure that I remain satisfied with it; and at any rate, it does not figure in *Kant on Reflection in Virtue*.

⁵ Kant's works are cited, as is customary, by volume and page of the German Academy of Sciences edition. The exception is the *Critique of Pure Reason*, which is customarily cited by the pagination of the first and second edition of the text, abbreviated A/B. The following other abbreviations are used: Anth=*Anthropology from a Pragmatic Point of View* (Ak. 7); MS=*Metaphysics of Morals* (Ak. 6); Rel=*Religion within the Boundaries of Reason Alone* (Ak. 6).

⁶ See *Kant on Reflection and Virtue* (29) and note 24 for references to Kant's works.

⁷ See *Kant on Reflection and Virtue* (29), and note 23 for further references to Kant's works.

But if this were correct, it would be hard to see how we could make sense of Kant's view that prejudice is *culpable*: we are responsible for it; and we ought to free ourselves from it. Part of what is going on here is, I think, a tendency to confuse the sources of prejudice from prejudice itself. Kant is *quite clear* that there is nothing culpable about custom, inclination, and imitation: these are just facts of human psychology, and nothing we could coherently aspire to free ourselves from. But prejudice is culpable, and this is why Kant's applied logic — and the discussion of reflection in this context — takes on a moral, even moralistic, cast at times. And what this shows, I think, is that we need my distinction between reflection-c and reflection-n to make sense of what is going on. Prejudice has to involve assent: that's why it is culpable. How do we understand this? I think Kant's apperception principle is a good guide: "The *I think* must be *able* to accompany all of my representations" (B131). Contrary to what McLear suggests, I do not misunderstand the modality of this. I do *not* think that Kant thinks that the "I think" is always actually accompanying all of my representations. The apperception principle says that a rational being is necessarily — constitutively, in virtue of its being the sort of mind that it is — able to consider its thoughts as its own. That sort of deliberate activity, whether undertaken on the occasion of a given judgment or undertaken in the adverbial mode I suggest later in the book, broadly falls under the heading of the normative conception of reflection. When you are thinking but not self-consciously accompanying these thoughts with the "I think," you are still moving from one thought to another in ways that show some sensitivity — even if imperfect — to cognitive norms like coherence. You aren't suffering random, unconnected states of consciousness; you are *thinking* one thing, which bears on how you can think other things, and so on. That involves having some kind of grip on yourself as the source of these thoughts as you think them, but it doesn't require that you are *doing* anything directly occupied with your own agency as a thinker. This grip on yourself, which comes online once you have come into the use of your reason, transforms what can figure as a mental state, and what it is like to think them. And if you don't have that at least tacit grip on yourself as a thinker, then I don't see how you can be thinking at all.

What then about taking things to be a certain way on the basis of prejudice? This is still directed by the shape of one's own mindedness. Prejudices are not random thoughts that are just induced in one by mechanical forces. They are the result of one's *taking* things to be a certain way — whether through custom, or the influence of prestige, or through inclination (habitual desire), and so on. Therefore, McLear overstates things when he claims that prejudice

and other sorts of judging of the first, merely content-presenting, variety is a *failure* to exercise one's cognitive agency altogether. This view does not make sense of the fact that we are epistemically responsible for prejudice, even though we are not meeting any number of epistemic standards or ideals when we allow our thought to be so driven. They are, indeed, an expression of our cognitive agency, but a corrupted one. Applied logic, which is concerned with this problem, is a therapeutic guide to the fallen human (epistemic) condition. And the sort of problems it takes up require the distinction between reflection-c and reflection-n.⁸

Finally, McLear raises doubts about my view that Kant takes reason to be a cognitive capacity through and through (and not just in its theoretical exercise); this claim is important for the transition from the basic cognitive virtue (which I associate with “healthy human understanding”) to Kant's conception of moral virtue in the second half of the book. In Chapter 4 of the book, I present the textual evidence — which I do regard as overwhelming — that Kant takes reason to be a cognitive capacity in its practical exercise. McLear replies that Kant never uses the phrase “praktisches Wissen,” although he acknowledges that there are plenty of mentions of practical cognition (*Erkenntnis*) and even some of practical wisdom (*Weisheit*). But nothing should ride on the former, he contends, since “cognition is not knowledge (e.g. there can be false cognition).” That is fine by my lights. Curiously, McLear does not seem to see *Weisheit* — wisdom — to be a kind of knowledge. I don't know how else to understand what “wisdom” could be, if not *some* kind of knowledge. Now, it may be that the non-wise person will have practical judgments of McLear's merely content-representing kind; however, I would say that there is normally tacit endorsement of what is thus represented, which is why passion, like prejudice, is culpable. Of course, we ought to be seeing things, and to be motivated, in the manner of someone who is wise, or has practical knowledge of the good. Kant, moreover, indicates in the Introduction to the *Metaphysics of Morals*' Doctrine of Virtue that he takes “*wisdom* in the strict sense, namely practical wisdom” — wisdom properly understood — to be nothing other than virtue itself (6:405).⁹

2.

⁸ One could say that reflection-c is the capacity that reflection-n exercises. (This seems to be what Sussman suggests in a footnote to his comments.) That would be fine, as long as we clarify that reflection-c transforms what it is to have mental states, and what they can be like. In the book (Chapter 3, pp. 107-110) I suggested that the distinction between them may ultimately be notional in the sense that the coming online of the one will always bring in its tow the coming online of the other; but the distinction between the two is required in the *account* of the requirement to reflect.

⁹ As he says this, he nods to the Stoic tradition, a point that will come up again in these remarks.

As it happens, that remark about virtue takes place in a stretch of the Doctrine of Virtue that I've more recently come to think of as "the Stoic interlude": it runs several pages (MS 6:405-409), includes a sympathetic allusion to the so-called Stoic paradoxes (only the sage is healthy, rich, free, and so on; non-sages are fools, who cannot genuinely be any of those things), discusses affect and passion at some length, and concludes by endorsing the Stoic principle of apathy as necessary for virtue. This brings me to Francey Russell's acute remarks about the affect-passion distinction. I think she is quite right in the problems she points to in my account of affect in particular; and my views have evolved somewhat as I have begun to work on a new project on Kant and Stoicism. So, I will begin by recapping what is quite correct in Russell's comments on this issue, then I will sketch how I now see the affect-passion distinction and its relation to reflection. Then, I will turn to some of the aspects of Russell's comments that are more directly related to the topic of moral virtue.

Maybe I should begin by expressing some reservations I now have about the fact that I effectively began the book with the discussion of affect and passion. I was trying to change the orientation of existing discussion on reflection from its narrow focus on logical reflection, i.e. the reflection that is involved in concept formation and which is a topic for pure general logic. So, part of my agenda was trying to get people to see that there is another notion of reflection that figures in *applied* general logic: i.e. reflection-n. But I was also trying to make the case that the topic of reflection is not just a concern for logic, whether pure or applied; that it has relevance for practical philosophy as well. Moreover, the way Kant distinguishes between affect and passion does — and I still think this — support my view that a distinction between constitutive and normative requirements to reflect is needed. But I now see that Kant's discussion of affect and passion draws quite directly from the psychological monism of the Socratic and Stoic tradition, and in its details, particularly from Seneca. My failure to appreciate this clearly when I wrote *Kant on Reflection and Virtue* largely accounts for my error in interpreting his notion of affect.

Let's start with my error. Now, Kant says that passion (*Leidenschaft*) involves reflection; indeed, he appears to distinguish affect from passion precisely on that ground. (I am going to use German terms here to avoid terminological confusion.) Since Kant speaks of *Leidenschaft* as a culpable disease — "a cancerous sore of practical reason" (Anth 7:266) and something that one brings on oneself — it was clear that it could not be meeting the normative standard of reflection. At the same time, *Leidenschaft* quite obviously involves taking a practical point of view on what matters, what is worth doing; and so, I took it to involve reflection in the constitutive sense. It is the expression of an evaluative point of view — only,

it turns out, a confused one. I still stand by that much of my story. The problem comes with affect (*Affekt*), which — with this set-up — then seems to be something that lacks evaluative or practical point of view altogether, since Kant characterises it as either making reflection impossible, or else just very difficult. He is vague. Rather than seeing the nuance here, I ran with the idea that affect is incompatible with reflection. But since *Leidenschaft* already lacks reflection in the normative sense, I took it that *Affekt* must lack reflection in the constitutive sense. I bit that bullet. Yet if that were right, affect would be something very extreme indeed. An example might be the sort of rage that leaves you shaking and disoriented afterwards. (I have experienced this once or twice; I hope you have not.) Russell suggests that *Affekt*, so understood, would only be at the extreme end — “something more like a trauma-level emotional experience.” I think that is right; and moreover, such an extreme interpretation of affect does not fit many of the *examples* Kant gives of affect in the *Anthropology*, as Russell indicates. It might perhaps fit extreme levels anxiety and fear; but it certainly cannot accommodate, for example, Kant’s inclusion of admiration (*Bewunderung*) as an affect. Admiration is the mode of our enjoyment of natural sublimity by Kant’s lights, and it involves a kind of astonishment in finding a wisdom in the order of nature that one did not expect to find (*Anthropology* 7:261): this is not obviously viable in someone who has *ex hypothesi* momentarily lost all grip on herself as a thinker and an agent.

Moreover, the extreme interpretation of *Affekt* as lacking reflection-c would be incompatible with psychological monism. Affects would be something that blows in from outside rational mindedness. The way that Russell puts this is to say that, on this reading, “much of our emotional lives would be oddly unexperienceable.” The way I was inclined to think of this in *Kant on Reflection and Virtue* was through an idea of *simplicity*: there’s just searing feeling that, at least for a time, overwhelms any capacity one might have to think even in terms of good and bad. There’s just some variation on extremely intense pleasure, or extremely intense pain. The problem here is to understand what Kant means when he says that they “more or less” lack reflection. As I said, I think the *examples* that Kant picks make clear that he can’t take this to mean that a sufferer of *Affekt* entirely lacks a point of view on how things are. If I am afraid, I see something as threatening, and so on. How should we understand *Affekt* then? I think this is a complex topic — and I think Seneca’s version of the Stoic theory of emotions or passions (*pathē*) is directly on Kant’s mind. I can’t go into those details here, but I don’t think there can be any loss of what I am suggesting is constitutive of rational mindedness, the typically tacit grip on oneself as the source of a point of view. What I can do here is at least point to the fact that Kant takes *Affekt* to be an expression of the faculty of

feeling, not the faculty of desire; and thus *Affekt* is not, simply as such, an engagement of one's agency.

Passions (*Leidenschaften*), on the other hand, are expressions of rational mindedness, and involve taking a practical point of view on what matters. They involve maxims. Glossing this, Russell suggests that *Leidenschaften* involves “the *ersatz exercise*” of the capacity to reflect, “that is a habitual and perverse misuse” of it. I think this is mostly right. What I think it might miss is that *Leidenschaften* involve substantive evaluative confusion, perhaps along lines akin to the Stoic view of the mental infirmity or illness underlying the *pathē*. Cicero presents the Stoic view of this issue, explaining that the mental infirmity at issue is ““a vigorous opining that some object is worthy of pursuit which is in fact not worthy of pursuit, that opinion being deeply attached and rooted in the mind”” (*Tusculan Disputations* 4.26).¹⁰ The *pathē* are expressions of assent to evaluative propositions — and in that sense are analysed as a kind of judgment. These judgments, according to the Stoics, are all false. They involve an evaluative confusion that is endemic to human mindedness. Substantively, they involve taking to be genuinely good what is not really good but may be preferable and worthy of selection for its value in making things go well for one. The root of the problem is an epistemic failure, namely our propensity to *opine*. Citing the old Stoic sources, he says: “‘opining’ is when a person judges that he knows something which he does not in fact know” (*Tusculan Disputations* 4.26). So opining is the expression of a default presumption of taking oneself to understand, or know, when in fact one is in no such position.

Now, this should ring a bell: for this is exactly what is *generally* at issue in prejudice by Kant's lights — namely, a default presumption that one is a knower, that one understands, when in fact one is in no such position. We might ask why this should be a *default*. I think the answer lies in the constitutive notion of rationality: that to be awake and thinking is to be taking a point of view on how things are and what is worth doing. So the default is that one takes oneself to see how things are and knows what is worth doing. This brings me to Russell's remark about why passions are difficult to correct: she suggests that this is because “the passionate person is to some degree rightly oriented: insofar as she takes what is merely subjective as if it were objectively valid, the passionate person displays some concern for meeting the standard of objective validity, hence her engaging in reflection and rationalisation.” I think it misconstrues things to say the passionate person is rightly oriented

¹⁰ For translation and commentary, see Margaret Graver, *Cicero on the Emotions: Tusculan Disputations 3 and 4* (Chicago: University of Chicago Press, 2002).

on these grounds; it is just an expression of default mindedness — of opining, if you will. But Russell is quite correct to note that the passionate person engages in (some sort of) reflection, and rationalisation.

On the topic of virtue, Russell raises one overarching issue, which — as I see it — has to do with the idea that virtue, as Kant evidently conceives of it, is a kind of struggle: “the moral disposition in conflict,” as he says in the *Critique of Practical Reason* (5:84). As Russell rightly points out, this might be read either as a comment about an individual’s moral life, or as a comment about our creaturely existence as human beings. As I see it, the second point is broadly at issue in Kant’s relation to the Stoic tradition. What Kant disputes in Stoic ethics is principally the idea that its sage is supposed to have transcended human nature: this is arguably how the impossibility of backsliding, once one has attained sagehood, is understood. Note that Kant accepts the idea that virtue, once gained, could not be lost: it is there when he endorses the so-called Stoic paradoxes in the *Doctrine of Virtue* (6:405), and I think it figures in his account of virtue in Part I of *Religion within the Bounds of Reason Alone* (6:47-8). More precisely what Kant says there is that the default human condition is to make the maxim of self-love the governing principle of one’s motivational economy, subordinating to it the moral law that is constitutive of reason in its practical cognitive employment. Becoming a good human being requires a “revolution,” which Kant characterises as “a single and *unalterable* decision” by which “a human being reverses the supreme ground of his maxims” — subordinating now the maxim of self-love to the moral law (Rel 6:48; my emphasis). But no human being can ever know that this unalterable decision has been made; and to suppose that it has is to be guilty of moral fantasizing and the worst forms of religious enthusiasm. Yet some such revolution must be recognised, Kant argues, as requisite for virtue.

Now, Russell asks: what is virtue? Is it moral perfection? Or does it lie somewhere in the struggle to be good — the struggle undertaken, gladly, from the fallen condition of humankind? Here is another way of thinking of Russell’s question: is virtue an ideal of perfection, or is it a post-revolution commitment to morality? It is not sufficient to say, as she does, that virtue is “the attitude in the struggle towards such perfection” — but I do think it is plausible to say, by Kant’s lights, that virtue is a joyous attitude proper to the *post-revolution* struggle towards perfection. That qualifier, as I see it, is essential. Of course, the matter of being post-revolution is noumenal: it is not open to experience, in any sense. But crucially, the struggle, it seems to me, is there in any event. If Kant does indeed take the moral law to be constitutive of reason in its practical exercise, then anyone has some dim grasp of moral requirement just as soon as she comes into the use of her reason. This is part of what Kant

conceives under the heading of the original predisposition to the good in human nature. He also thinks that each will by default pursue her own happiness and tend to regard it as objectively or unconditionally *good*: and this is the fundamental evaluative confusion that, as I see it, Kant takes to be endemic to human nature. So anyone should have experienced some sort of struggle between the morality and self-interest. You will experience that conflict whether or not you have undergone a revolution in your soul, because the moral law is constitutive of reason in its practical exercise. So I think *struggle towards* perfection is probably not enough to capture the idea of virtue, unless you add that this is undertaken joyously in a post-revolution frame of mind. At any rate, I am inclined to think that virtue is an ideal by Kant's lights, something which we can only conceive, in pure thought, through the moral law. We can form images of ways in which this ideal might take shape concretely, in people who are situated and minded and temperamentally disposed in perhaps a wide range of ways. But those sketches are, perhaps unavoidably, always somewhat impressionistic.

One final remark about Kant's conception of virtue and ancient sources. I am somewhat uncomfortable with the idea that my account of Kantian virtue is Aristotelian, which is suggested by both Russell and Sussman. Although some great work has been done developing a Kantian account of virtue along Aristotelian lines — perhaps above all by Barbara Herman — as a historical point, I do not think there is much reason to think that Kant ever thought that long or hard about Aristotle, or even that Kant's fundamental systematic commitments align all that well with Aristotle's. As I hope to show in my next project over the next few years, there is ample evidence that Kant thought long and hard about Stoic ethics, particularly through Seneca, and that he endorses the psychological monism of the Stoics, as well as something along the lines of their cognitivism about virtue. Russell rightly points out that my view of virtue bears some resemblance to John McDowell's account in "Virtue and Reason" (1979): McDowell was undoubtedly, and in ways that I probably can't even quite appreciate, an influence. However, I am also inclined to think that McDowell may have mischaracterised his own view of virtue by calling it Aristotelian. By my lights, McDowell's view of virtue belongs more to the Socratic and Stoic monistic tradition — but that is not a claim I could elaborate here.¹¹

¹¹ I haven't said anything here about Russell's interesting proposal that my argument that attention is required for "full blown experience" should imply, more or less by the same token, that "the capacity for *aesthetic* experience is also required for full-blown experience." I am inclined to think that the connection that Russell wants to draw here would require some kind of view about how Kant's conception of experience relates to his distinction between theoretical and practical cognition — which is a difficult topic, and not one I directly took up in the book. But I would like to see how the idea might be developed.

3.

David Sussman focuses on the final parts of my book, about moral virtue. He expresses my view quite accurately, and certainly more elegantly than I do. He also says that the case I make for the essentially cognitive nature of Kant's conception of moral virtue, while having "considerable truth," is nevertheless somewhat overstated. That may be: I was trying to reorient the discussion. So, in reply to Sussman, I want to elaborate on how I understand some of the issues that arise here about the cognitive nature of human moral psychology by Kant's lights. This will be a development of some of what already came up in response to Russell's comments, particularly on the topic of the passions.

Sussman gets the fundamental claim about my view of virtue exactly right, when he says that a virtuous person "is someone who has a vivid kind of appreciation of the central moral value (essentially, of just what a person is), and [is] able to see how this value is at play in the various features of the particular circumstances we are faced with." The way I see it, our grasp of the fact that "this is a person; not a thing" admits of degree — of clarity, of thickness, and of general practical determinacy. The virtuous person may not be able to express, propositionally, the knowledge she has; rather, as Sussman so nicely glosses my view, "the appreciation or attunement [at issue] is cognitive in the sense that having an ear for music might be, or a sense of humour, or a feel for the strength of a position in chess." While Sussman thinks that much of the picture is right, he also thinks that moral virtue "also has distinct volitional elements, at least insofar as it makes sense to still talk of any contrast between the cognitive and the volitional at all."

To elaborate, Sussman points to Kant's conception of virtue as a kind of strength — summed up in his Latin gloss *fortitudo moralis* (MS 6:380) — and he notes that Kant "tells us that virtue is a matter of having a moral resolve that is powerful enough to overcome whatever obstacles that inclination puts in our path." Later on, he acknowledges a point that I stress in the book: namely, that there is nothing *wrong* with the inclinations per se. Kant's exact words about virtue as strength are as follows: "strength is required, in a degree which we can assess only by the magnitude of the obstacles that the human being himself furnishes *through his inclinations*" (MS 6:405; my emphasis). (This passage, too, comes from the Stoic interlude in the Doctrine of Virtue.) Now, Kant does not here claim that inclinations themselves are the obstacles; and Sussman, I think, recognises this when he notes that the obstacles are not the inclinations but rather what we do "in response" to them. Kant's point, as I see it, is that the

obstacles to virtue are the *passions*; and I want to elaborate briefly on this point in the hopes of making more plausible my cognitivist conception of Kantian virtue.

Sussman says that “[t]his feature of virtue” — namely, the business of its strength to overcome obstacles furnished through the inclinations — “doesn’t sound very cognitive.” He goes on to give examples of clear-eyed weakness of will. Here are two of Sussman’s examples: You know full well that eating pizza at this hour will give heartburn; but you do it anyway. You know full well that you should not commit adultery; but you do it anyway. I’m not entirely sure what to say about the pizza case. Doctrinally I am a committed psychological monist, so I would be inclined to explain that case as exhibiting shifting and unstable views about what is good or at least worth going for. But I want to focus on the cases where moral value is more immediately at issue, like the adultery case. Of course, inclinations seem to come to us unbidden; and some may be quite unwelcome. But there are *many ways* of “knowing that adultery is wrong.” Suppose you find yourself with an inclination to flirt with someone who is not your spouse; suppose that you have this inclination *a lot*, so that thoughts, fantasies, perhaps even plans of adultery get involved as well. In what sense do you know that “adultery is wrong”? Let us ask what you do in response to these inclinations. Do you gratify them, even if not in deed, but in rumination? You might still know that adultery is something you should not do, but I do not think such knowledge is very robust. But perhaps you stop and look around you, look your spouse in the eye when you talk about what to have for dinner or who will pick up the kids: *Here is the person that I would hurt* by carrying on with this. Maybe then you think, in rich ways, of your past history, the commitments, the life shared. Intuitively, such a person might know that adultery is wrong in a somewhat different way than the first. Such knowledge might not be compatible with the gratification of this inclination, even just by ruminating on it.¹²

¹² I don’t think the passage that Sussman quotes from the *Religion*, where Kant alludes to the Apostle Paul from Romans (7:14-15) entails that Kant endorses clear-eyed weakness of will, and in turn a fundamental separation between motivational and cognitive aspects of practical reason. But I grant that the passage is challenging for my reading. Interestingly, what Paul actually says right before Kant begins his quotation is “I do not understand what I do [ὁ γὰρ κατεργάζομαι οὐ γινώσκω].” Paul then says: “what I want, this I do not do, but what I hate this I do [οὐ γὰρ ὁ θέλω τοῦτο πράσσω, ἀλλ’ ὁ μισῶ τοῦτο ποιῶ]” (Romans 7:15). Kant renders it, very loosely, thus: “Wollen habe ich wohl, aber das Vollbringen fehlt” (6:29). Kant then elaborates with the following explanation: the good that I will, and have incorporated into my maxim, *would be* “an irresistible incentive objectively or ideally (*in thesi*)” but subjectively — “*in hypothesis*” or under the human default condition — is “the weaker (in comparison with inclination) whenever the maxim is to be followed” (Rel 6:29). If I were a holy will, my grasp of this good would be perfect. But I am not a holy will, which is to say that I am subject to fundamental evaluative confusion that is endemic to human beings. I have other views of what would be good, which acquire considerable force through the accretions of my own assent. They can overpower my understanding, which is weak and indeterminate, of what really is good, of what I can recognise as moral requirement.

Passions (*Leidenschaften*) involve maxims, Kant says. They are an expression of rational mindedness, of taking a point of view on what matters, what is worth doing for what, and so on. Kant explains passion also as a kind of inclination, which he regularly glosses as “habitual desire”:¹³ they have the force of the habitual accretions of one’s own assent, over and over again, to their principle. This is how Kant conceives of passions as self-wrought, but also in their way overpowering: they work like water that digs itself deeper and deeper into its bed, as Kant memorably puts it (Anth 7:252). They are also culpable, and not for some silly reason that it is in itself bad to be habitual about anything, especially desires. They are culpable because of the fundamental evaluative confusion that underlies them. The person who supposedly knows full well that adultery is wrong and yet finds herself carrying along a track leading to it, has something in her that sees this flirtation as good, as to be pursued. What those thoughts might involve will be particular in ways too various to try to elaborate from the armchair. Schematically, it involves assent to a practical principle, a maxim. This kind of direction of one’s thought, I am suggesting, could be understood as a kind of confusion about what really matters, and what really is good. Put more simply, we should doubt that she really “knows full well” that adultery is wrong.

I’d like to end with the issue about self-conceit, and delusion about one’s own moral worth, that I think both Russell and Sussman, each in different ways, have raised. As Sussman puts it, it is owing to our “radical evil” that “we endlessly rationalise and deceive ourselves about what we are doing and why.” That’s right; and as I see it, that sort of point is best understood with the kind of cognitivist, monist psychological framework that Kant develops from the Socratic and particularly Stoic tradition. Elaborating, Sussman says: “It is not just that our vision is often blurry; but rather, that we see most clearly when we are in the grip of an illusion.” What he means by this, I think, is that mongering for reasons and justifications can just as well feed our self-conceit, and moral fantasies, as anything else. I quite agree. This is one reason why I wanted to show the roots of Kant’s conception of virtue in a standard of healthy human understanding; it is also why I urged, with Iris Murdoch somewhat secretly in mind, that appreciating the reality of other persons — really, deeply, thickly getting that *here is a person*, and not a thing — is perhaps what most properly halts the smooth thrum of such self-justifying trains of thought. Of course, when a plausibly conscientious person realises that

¹³ The German is *habituelle Begierde*: see e.g. Anth 7:251, 265; MS 6:212; Rel 6:28. For discussion of the point, see *Kant on Reflection and Virtue* (39-40).

her understanding of this fundamental point has been deficient here, now, with regard to this person, then she learns something about herself. Probably something hellish, too, if she is a plausibly conscientious person. Note that Kant uses the same language for his presentation of the three maxims in the *Anthropology* (7:228) as he uses in his presentation of the Delphic command in the Doctrine of Virtue (MS 6:445): both are said to be, effectively, the beginning of human wisdom. Together, Kant is perhaps making a point that moral self-knowledge is possible only through the outward engagement of our thought, where concrete others play a special role. Sussman is right to point out that Kant's account of virtue is Christian: at its centre lies an agapic conception of love, a welcoming being-pleased in the existence of others. This love is an ideal. I also think that, for Kant, it might be the best and truest way of *knowing* the value of persons.

Bibliography

Kant's works are cited by volume and page of the German Academy edition, with the exception of the *Critique of Pure Reason*, which is cited according to the pagination of the first and second editions, abbreviated A/B. (See note 5 for other title abbreviations used.) Quotations are drawn from the translations in the Cambridge Edition of the Works of Immanuel Kant, series editors Paul Guyer and Allen Wood.

Allison, Henry. *Kant's Theory of Taste*. Cambridge: Cambridge University Press, 2001.

Graver, Margaret. *Cicero on the Emotions: Tusculan Disputations 3 and 4*. Translation and commentary. Chicago: University of Chicago Press, 2002.

Longuenesse, Béatrice. *Kant and the Capacity to Judge*. Translated by Charles T. Wolfe. Princeton: Princeton University Press.

Lu-Adler, Huaping. *Kant and the Science of Logic*. New York: Oxford University Press, 2018.

McDowell, John. "Virtue and Reason." *The Monist* 62, no. 3 (1979): 331-350.

Merritt, Melissa. "Varieties of Reflection in Kant's Logic." *British Journal for the History of Philosophy* 23, no. 3: 478-501.