

Melissa Merritt, *Kant on Reflection and Virtue*.
Cambridge: Cambridge University Press, 2018. 219 pp.

In *Kant on Reflection and Virtue*, Melissa Merritt offers a detailed interpretation of what she calls “the Kantian reflective ideal,” bringing together Kant’s most important claims concerning the value of reflection and our putative obligation to reflect. At its most minimal, Merritt takes the ideal to be implicit in Kant’s claims that every judgment requires reflection (*Überlegung*) (A/B: 260–61/316–17), and that such reflection is a duty (A/B: 263/319).¹ But for Merritt the reflective ideal goes beyond a supposed duty to reflect. It also points to the “supreme value Kant accords to being rationally reflective” (2). The ideal therefore covers both the axiological and deontological theses Kant advances with respect to reflection and its place in human life.

Merritt’s aim is to offer a novel interpretation of this ideal that will help dispel its caricatured misinterpretation. The widely accepted caricature presents reflection not only as impossibly demanding but also “precious, hyper-deliberate and repugnantly moralistic” (2), requiring agents to step back from each and every judgment so to scrutinize its grounds. On Merritt’s account, Kant’s ideal reflective agent judges “in the right spirit” or “right frame of mind,” where this ultimately means following the three maxims of the healthy understanding (52): ‘to think for oneself’; ‘to think in the position of everyone else’; and ‘always to think consistently’. One who has cultivated their cognitive capacities in conformity with these maxims has, according to Merritt’s novel reconstruction, developed general cognitive virtue or, equivalently, good cognitive character. While Kant has no explicit conception of cognitive virtue as such, Merritt argues that such a conception is implied by the principles of healthy understanding, which specify a categorically required way of thinking for beings like us. By developing good cognitive character in accordance with these principles, argues Merritt, we exercise the self-determination proper to a rational being in the most basic sense. This self-determination does not involve our stepping back from every judgment to scrutinize its grounds; rather, good cognitive character just is the practical know-how we express whenever we make good use of our cognitive capacities at all.

In part I of the book, Merritt lays the groundwork for her account by distinguishing between normative and constitutive requirements to reflect, which are often muddled together in Kant’s own discussions and in secondary literature. This clarifies the normative reflection (reflection-n) at issue in the

1. The works of Immanuel Kant are cited according to the standard citation practice associated with the Akademie edition of his writings.

reflective ideal. Merritt then goes on to show that reflection ought to be understood in terms of the three maxims of healthy understanding. In the book's second part, Merritt argues for what she calls the "specification thesis," the claim that all virtue is cognitive virtue, which is nothing other than Kant's notion of healthy understanding, and that moral virtue should be understood as a specification thereof. This ambitious thesis casts the three maxims of healthy understanding as fundamental principles guiding the good use of reason in all its employments, both theoretical and practical. As Merritt is well aware, this thesis presupposes a controversial conception of the unity of reason according to which reason is at bottom a cognitive capacity.

The book's final part explores Kant's conception of virtue from this new angle, elaborating on the cognitive basis of moral virtue. Merritt attempts to show what we gain by considering moral virtue a specification of general cognitive virtue, understood as healthy understanding. Her discussion builds on a novel reading of Kant's conception of virtue that Merritt calls the "skill thesis," namely, the claim that "moral virtue is a certain sort of free skill, one governed by the adoption of morally obligatory, rather than discretionary ends" (184). Skill, according to Merritt, is "practical intelligence" "concretely embedded as a disposition for action" (184). A skill is 'free' when it involves consciousness principles, like those of healthy understanding. The upshot of Merritt's argument is that reflection is just the practical intelligence that allows us to realize healthy understanding *in concreto*; such reflection is something we categorically ought to care about and pursue.

As this précis suggests, the volume has many strengths that recommend it to anyone interested in Kant. Giving the principles of common understanding a bedrock normative role in critical philosophy provides a fresh take on questions of growing scholarly interest, especially those bearing on epistemic normativity and the status of doxastic maxims. A further merit of the book is its insistence on finding unity in Kant's account of reflection, which is scattered and frequently thematized under the auspices of other concepts. This linking together of diverse areas of discussion is extremely helpful. The general account of virtue that Merritt develops and attributes to Kant, however controversial exegetically, is philosophically rich, and worth studying apart from whatever it may tell us about Kant.

In what follows, I wish to develop two critical points that, I hope, will illuminate the book's central claims. Although Merritt's account of virtue as good cognitive character is compelling, it is also ambiguous in an instructive way, pointing to unresolved questions of a more fundamental kind. On the one hand, Merritt argues that what makes the good use of our cognitive capacities good is that this use expresses the "the right frame of mind," or is made "in the right spirit" (54). These adverbial characterizations are then fleshed out in different ways, not all of them obviously equivalent. Early in the book, for instance, Merritt characterizes the reflective person (the person who makes good

use of her cognitive capacities) as someone who takes a proper interest in those capacities. On this account, it appears to be the interest motivating the use of our cognitive capacities that makes that use good. This interest is cast in resolutely inward-looking terms: the reflective person does not take an interest in the *object* of the good use of her capacities (whether truth, knowledge, or some other end) but rather in the capacities themselves. Later on, however, Merritt's characterization shifts. In chapter 5, for instance, she argues that healthy understanding "calls for the engagement of one's cognitive capacities from a particular interest, which should be conceived as respect for truth" (151): "There is a purpose that is internal to any exercise of cognitive capacities at all—knowing—and to adopt this end freely is to respect truth, which itself calls for developing the strength to submit to it. Respect for truth is therefore a proper part of the resource at issue. And this is finally what makes healthy understanding a natural perfection like no other" (156).

Taking an interest in one's cognitive capacities, in knowing, and in respecting the truth might all, finally, be different expressions of the same fundamental interest. This is not obvious, however, and requires more argument. I draw attention to this issue because it bears on the more substantive and genuinely difficult question of whether there is a fundamental value or end (and a corresponding interest in that value or end) that justifies and/or makes good all uses of reason in the first place.

As we have seen, Merritt argues that there are fundamental principles in virtue of which any use of our cognitive capacities, practical or theoretical, counts as good: the principles of healthy understanding. She also argues that conformity with these principles must flow from a certain interest. Nevertheless, she appears undecided whether an obligatory end or fundamental value plays a justificatory role in explaining the above. One might argue, as Kant does, that categorically required principles specify obligatory ends, such that one cannot affirm the legitimacy of such principles without thereby affirming the ends they specify. Merritt herself seems to concede this, writing that "healthy understanding" "is a way of thinking that answers to some highly general end of knowing, which is obligatory. One is not at (moral) liberty to reject this end. The telic drive internal to our cognitive capacities is knowledge, the desire to know is constitutive to the use of these capacities" (143). Here Merritt presents knowing as an obligatory end, and at least implies that using cognitive capacities correctly involves seeking to realize, or taking an appropriate interest in realizing, this end (via the principles of the healthy understanding). This sounds plausible enough as an account of theoretical cognitive capacities, but it is unclear how knowing could function as the obligatory end of all uses of reason, given that its practical use is, on Kant's account, aimed not at knowing but at the highest good.

In claiming that all uses of reason are fundamentally cognitive, Merritt seems to want to contest this received view of the divergent ends of practical and

theoretical reason. She argues that Kant's repeated reference to "practical cognition" should be taken to secure this claim, where cognition refers fundamentally to a "mode of knowing" (115). She notes judiciously that practical and theoretical modes of knowing are indeed very different: "practical rational cognition does not merely determine its object 'but also makes it actual'" (115). As "the efficacious determination of the will by rational principles" (117), practical cognition is concerned with our capacity to cognize the good, not objects of experience, through the determination of the will.

Despite these allowances, however, Merritt's cognitive conception of the practical use of our rational capacities is unpersuasive, because the sense in which the determination of the will is to count as a "mode of knowing" remains unclear. Merritt attempts to address this problem by claiming that "the distinction Kant draws between cognition and volition is drawn under the umbrella of the idea that reason is at bottom a cognitive capacity" (118). This highly contentious claim is hard to interpret and requires more argument than Merritt provides. Does Merritt's view imply that the will itself is not a causal power but a cognitive one? Can causal powers be fundamentally cognitive on Kant's account? Without fuller answers to these questions, it is difficult to understand and endorse Merritt's contentious claim that reason as such is fundamentally a cognitive power.

This difficulty returns us to the earlier worry about fundamental ends. There appears to be a real discrepancy between exercises of virtue aimed at knowing and expressing a respect for truth, and those aiming at cultivating the will and expressing a commitment to the fundamental moral value, which on Merritt's account is that of humanity. To argue, as Merritt does, that you cannot exercise the latter type of virtue without also exercising the former (but not vice versa) does not fully settle the normative question about which ends and values are fundamental and why. Merritt addresses this issue by appealing to Kant's distinction between the original and highest *Bestimmung*, or vocation, of the human being. While our original vocation consists in the development of healthy understanding, our second and higher vocation consists in the cultivation of the will. Merritt writes, "while the cultivation of our cognitive capacities is fundamental to this 'original' *Bestimmung*—we cannot realize our rationality without it—the end for which this is all done is ultimately the cultivation of the will, or practical reason" (123). That all of reason has a shared cognitive basis and that all good uses of reason answer to the same principles should not be confused, she argues, "with any claim about what is the most important exercise of reason, or what is its highest calling. . . . The highest calling . . . consists in the cultivation of specifically practical cognitive capacities" (123).

This attempt to establish a hierarchy in the uses of our rational capacities is promising and worth developing further. In the absence of this fuller development, however, a tension lingers in Merritt's account between rational ends that seem more naturally suited to the good theoretical use of our cognitive

capacities, and ends that seem more naturally suited to the good practical use of those capacities. This tension is brought out strikingly in the book's final paragraph:

The guiding thread of this book is that Kant conceives of reflection in resolutely cognitivist terms. It is in the interest of knowing, and that alone, that we ought to be reflective. The motto of the Kantian reflective ideal is *self-determination through understanding*. There is an element of submission to this, of owing fidelity to what obtains independently of any particular exercise of cognitive capacities. This includes fidelity to a value that obtains independently of anything anyone does or doesn't do, cares about or doesn't care about. This submission, moreover, is not to any indifferent fact about what is the case. We owe fidelity, chiefly, to what is most worth caring about in the complete order of things. For Kant, this can only be actual human beings. (208)

In Merritt's final accounting, the value of knowing appears derivative; the value of actual human beings, fundamental. One wonders whether this ordering of values is fully compatible with the primacy Merritt accords to the principles of the understanding as the ground of virtue as such.

Sasha Mudd

Pontificia Universidad Católica de Chile

Philosophical Review, Vol. 130, No. 2, 2021

DOI 10.1215/00318108-8809958

Christine M. Korsgaard, *Fellow Creatures: Our Obligations to the Other Animals*. New York: Oxford University Press, 2018. xiii + 252 pp.

Over the last several decades, Christine Korsgaard has produced a tightly woven body of moral philosophy that incorporates strands from philosophy of mind and history of philosophy as well. *Fellow Creatures* deepens and extends this body of work and also provides an overview of her previous work. This book is essential reading for anyone interested in Korsgaard's philosophy or anyone who is concerned about the ethics of our treatment of animals.

Fellow Creatures is a strikingly honest engagement with a broad range of work in animal ethics and the problems that this work addresses. Korsgaard's

Many thanks to Alice Crary and Beatrice Longuenesse, whose comments on an earlier draft provoked some significant improvements.