Kantian Review (2022), **27**, 469–474
doi:10.1017/S136941542200019X

KANTIAN REVIEW
CAMBRIDGE UNIVERSITY PRESS

AUTHOR MEETS CRITICS

Comments on Katharina T. Kraus, *Kant on Self-Knowledge and Self-Formation: The Nature of Inner Experience*

Allen Wood

Indiana University, Bloomington, IN, USA and Stanford University, Stanford, CA, USA
Emails: awwood@indiana.edu, allenw@stanford.edu

Abstract

Kraus's book is both deep and wide-ranging. My comments focus on her account of Kant on self-awareness – both *a priori* and empirical apperception. Basic to her account is what she calls the hylomorphism of mental faculties in Kant. Kraus distinguishes her 'reflexive' account of apperception from both 'logical' and 'psychological' accounts. An inevitable question is: Does Kant think we have an empirical *cognition* of the self? Kraus seems to want to say yes, but I question this answer. Cognition requires both intuition and conception. My claim is that it requires intuition in both space and time, but inner empirical self-awareness is apparently in time only. Kant's Refutation of Idealism in B, as developed later in the Kiesewetter essays, makes awareness of our body essential to time determination.

Keywords: self-awareness; apperception; cognition; intuition

Katharina Kraus's new book is an impressive, wide-ranging treatment of Kant's theory of self-consciousness and selfhood. It is well-informed on the literature and penetrating and original in its reading of the Kantian texts. The book covers too much to make possible an informative but comprehensive set of comments so my remarks will have to be selective.

The book has two basic themes: self-knowledge and self-formation. The first deals with theoretical issues concerning Kant's account of self-awareness or apperception, both transcendental and empirical, and its role in the Kantian account of experience or cognition. The second theme is practical: how to be (and become) a self on Kant's account. In my judgement, the book does a better job with the first theme, which occupies most of it. The second theme is also harder to treat, since it threatens to expand into a complete account of both prudential reason (pragmatic) and moral reason (or virtue). Given my own interests in Kant, I expected to read the book for the second theme, but I think I learned more, and perhaps also found more to discuss, in Kraus's treatment of the first theme.

For me the heart of the book was found in chapters 3 and 4, which deal with apperception and both transcendental and empirical self-awareness. This is well-trodden territory in the Kant literature, but also a difficult and complex topic, rich enough

for further (perhaps endless) discussion. It is an understatement to say that it is controversial, and also that it is extremely difficult to put together a clear and self-consistent account of it from the things Kant says even in those parts of the first *Critique* that deal with it. They would have to include, at a minimum, the Transcendental Deduction of the Categories (especially the B edition version) and the Paralogisms (both in A and B). I think Kraus's account is not always compelling based on the texts, but its strength lies in the insight she displays in seeing how texts that are familiar to all students of the *Critique* can be read with subtlety and discernment in new and interesting ways. The account that emerges provides Kant with original answers to questions most readers (I speak of and for myself here) have not even known how to ask before reading her book. When I say I do not always find her reading textually compelling that should not necessarily be understood as a criticism. It might more easily be read as a positive assessment of it. For I confess that she has turned me into a learner and an inquirer into the topics she is discussing. It is rather that she questions the ways I have read certain texts for many years, and suggests readings of some of them that I have never considered, much less accepted or rejected. It may be that longer and deeper reflection will convince me she is right about things that I have been unable to take in fully on my first reading of her book.

Basic to Kraus's account is what she calls the hylomorphism of our mental faculties. This would seem a product of the Kantian idea that experience results from the application of what those faculties contribute – which is form – to what is given through sense-experience – which is matter. This corresponds to the Kantian distinction between what is transcendental and what is empirical. She distinguishes within Kant's theory of representation in general between two crucial and complementary aspects: reflexivity and referentiality. Every representation involves a relation to its subject, but also refers to what is, in the broadest possible sense, its object. This distinction is different from but closely related to another, which seems to me fundamental to Kant's theory of experience: that between perspectivity and objectivity. Every experience and every cognition is from a distinct point of view – in both space and time, and belonging to a unique individual self; but every cognition is cognition of an object, making judgements about it which if true are universally valid for all subjects. This last distinction seems to me to ground Kant's duality between intuition, which makes cognition perspectival, and understanding, whose judgements give it universal validity. Even subjective representations, however, have both reflexivity and objectivity in a broad sense, even if (as with sensations and feelings) their objects are present only to this subject and have no objective validity.

What is apperception? Kant says many things about it, and it plays an important role both in the Transcendental Deduction of the Categories and in Kant's critique of the psychological paralogisms. Clearly apperception is self-awareness. But as Kraus emphasizes, Kant distinguishes pure or transcendental apperception from empirical apperception. And Kant says many things about apperception, linking it in subtle ways to our conscious states and ways of being aware of them as well as of oneself as the subject of them. Kraus is painstaking in her treatment of apperception, especially transcendental apperception. She considers the very large secondary literature on this topic, and tries to present an account that enables her to relate her own account to the main alternatives found in the literature, and to offer a critique of them.

Kraus takes apperception to be closely related to the 'I think' that must be able to accompany any representation, as Kant famously says in §16 of the B Deduction. Before

https://doi.org/10.1017/S136941542200019X Published online by Cambridge University Press

offering her own account of apperception, Kraus considers two other accounts, which she calls the 'Psychological' account and the 'Logical' account. The psychological account takes apperception to be awareness of a mental act – specifically, the act of synthesizing the manifold. The logical account takes apperception to be the self-ascription of a mental state, or perhaps the capacity for its self-ascription. Although Kraus cites various interpreters in presenting these alternatives, I don't think she wants to claim that Kant scholars fall neatly into two groups under these headings. For she seems to realize that many accounts are hybrids to one degree or another. I also don't think Kraus means to assert that either account that she rejects is being rejected because it is simply *wrong*. For it is clear that Kant does think that apperception is an *Actus* of the intellect, and involves our awareness of 'self-activity' (as Kant puts it at B130), even the self-activity of combination or synthesis. And it would be paradoxical beyond imagining to try to deny that 'I think R' involves the self-ascription of R to the subject I: for that's just what the words 'I think' explicitly do. At most, Kraus is denying (and with good reason) that apperception is a mental state whose *object* is the act of synthesis; for it seems to be a pre-reflective act that *does not* have itself as an object. And since apperception consists not in any actual self-ascription but rather the permanent possibility or availability of the 'I think', she is also right to deny that it consists in an actual self-ascription of any particular representation. Kraus even prefers to say (as she does about the 'logical' reading of apperception) not that it is false but that it is 'inadequate' in that it does not account for the transcendental function of apperception in Kant's theory of consciousness and self-consciousness.

Kraus's preferred account of apperception is what she calls the 'Reflexive' account. By this she seems to mean that it captures the pre-reflective self-awareness that belongs to any conscious representation or mental state. As such it constitutes, following the Kantian hylomorphism, the *form* of every representation simply as such, whatever its content might be. This seems both correct and insightful, and even echoes verbatim what Kant says in the Paralogisms about 'the simple and in content for itself wholly empty representation **I** . . . [which] consciousness in itself is not even a representation distinguishing a particular object, but rather a form of representation in general, insofar as it is to be called a cognition; for of it alone can I say that through it I think anything' (A346/B404).

Apperception, the form of reflexivity and the synthetic activity it thereby indicates, does not apply only to objects in the strong sense relating to cognitive judgements which Kant introduces in §§17–18 of the B Deduction. Kraus makes the point that Kant's discussion of combination and apperception in §§15–16 applies more broadly than to cognitive judgements, or 'judgements of experience' in the terminology of the *Prolegomena*, so that it pertains equally to 'judgements of perception', to desires, aesthetic judgements and even to subjective feelings. For these too have the form of reflexivity involving apperception, synthesis and the availability of the 'I think'.

Another striking feature of Kraus's presentation of the apperceptive 'I think' is her claim that it should be understood as 'expressive'. If I say 'This rose is red', I am *asserting* something about the rose and its colour. But I am also thereby *expressing* (without directly *asserting*) that *I believe* (and am disposed to assent to) the judgement that the rose is red. Expression, in this sense, is a performative or illocutionary act, in the sense used by Austin or Searle. And as Kraus points out, it corresponds to some ways Kant himself uses the term *Ausdruck* and the verb *ausdrücken*. The apperceptive

https://doi.org/10.1017/S136941542200019X Published online by Cambridge University Press

'I think' *expresses* 'the subject realizing the original synthetic unity that is necessarily required for a significant representation, without determining anything about the subject' (p. 123). That, I take it, is why Kraus wants to deny the 'psychological' reading, if it takes apperception to be some kind of judgement *about* the synthetic act of the subject, and also the 'logical' reading, if it takes the 'I think' to involve a judgement self-ascribing the representation to the subject.

A bit ago I quoted a passage in which Kant said of the 'wholly empty representation **I**' that it is 'a form of representation in general, insofar as it is to be called a cognition'. In this rather coy statement, Kant raises an important question we might have about his theory of self-consciousness, and which I also have about Kraus's account of it. Is our inner self-awareness for Kant, as it is found in apperception and even in empirical self-consciousness of inner sense, a self-*cognition*? The passage just quoted, through its use of the term *insofern,* seems to say that it is, but also to suggest that any such claim has to be significantly qualified in some way.

Kraus makes it clear that the I (or self) cognized in inner sense is not a substance. But it is also not a mere 'bundle' of representations. In chapters 5–6, she presents an elaborate theory of self-cognition based on the regulative idea of the soul which is less than a noumenon but more than a mere heuristic fiction. Kraus attributes to Kant the epistemic project of treating the subject of inner sense as if (*als ob*) it were a persisting substance, and a 'mental whole' that is an object of inner self-knowledge – a theory of which she develops in chapter 6. There are certainly strands of Kantian text that support this.

But I have a worry about it nonetheless. For Kant, cognition (*Erkenntnis*) requires both intuition and conception. A cognition is a representation which is referred to an object and brought under a concept (A50/B74). But for beings like us, any cognitive relation to an object must, either directly or indirectly, refer to intuition; and our intuition is always sensible (A19/B33). We have two forms of sensibility: outer sense (space) and inner sense (time). The former intuits what is different from me, the latter my own states.

Is inner intuition by itself sufficient to make cognition of an object possible? Or must any object, in order to be cognized by us, be intuited both in inner sense and outer sense? I did not find in Kraus a clear, consistent and unambiguous answer to this question. She seems aware of the tension, and approaches the question here and there, but if she has a single decisive answer, I missed it. Perhaps there is no such answer in Kant either. One answer that might be ascribed to Kant is that yes, we do cognize ourselves as objects of inner sense, but not as we are but only as we appear to ourselves (B155–6, *Anth*, 7: 142). That is what he says, or seems to say, at several points. But it is not clear what it means.

It could mean just that we cognize ourselves, as we do the rest of nature, only as appearances, not things in themselves, or only as phenomena, not as noumena – whatever those claims mean (also a matter of much dispute). But perhaps Kant thinks our self-cognition is more limited or qualified even than that of the rest of nature. He often suggests that human self-knowledge is always afflicted with dissimulation (*Anth*, 7: 121, 132–3, 332; *G*, 4: 408; *MS,* 6: 441, 447) so that the claim that we cognize ourselves only as we appear to ourselves has a stronger meaning in the case of self-cognition than it does applied to the rest of nature. The proposition that I cognize myself only as I appear can, Kant says, be 'maliciously twisted so as to mean: it only seems to me that I have certain representations and sensations, or even that I exist at all' (*Anth*, 7: 142). But Kant himself seems at times to be sympathetic to such a malicious interpretation of what people

https://doi.org/10.1017/S136941542200019X Published online by Cambridge University Press

believe about themselves, so he may not be entirely repudiating it when he says this. Nietzsche claims that our inner lives are even more phenomenal, even more merely apparent, than the outer world: 'The phenomenalism of the inner world. Nothing is so much deception as this inner world which we observe with the famous "inner sense."' (*Will to Power*, trans. Walter Kaufmann and R. J. Hollingdale (New York: Random House, 1968), §§476–9). Is it wrong to suggest that Kant anticipated Nietzsche here? I don't think it is. Nietzsche could even have been quoting Kant.

Kant does reject decisively (as absurd or even silly) the suggestion that we only seem to ourselves to exist. (Think of his derisive reference in *The End of All Things* to mystical Taoist philosophers 'sitting in a dark room with their eyes closed exerting themselves to think and sense their own nothingness' (8: 335).) But Kant does take more seriously the thought that inner experience alone cannot afford genuine cognition of an object: 'Inner sense sees the relations of its determinations only in time, hence in flux, where the persistence of what is observed, which is necessary for experience, does not occur' (*Anth*, 7: 134). The fact that inner experience is only in time, not in space, generates worries for Kant when it comes to our belief in our personal identity through time as an object of inner intuition (A365, B420). In the B edition Refutation of Idealism, Kant argues that the determination of our existence in time itself does not depend on the mere representation of objects in space but in our cognition of them as actual (B275–9). The objective cognition of our existence as determined in time thus depends on outer intuition of objects in space. It has always seemed to me, therefore, that this requires Kant to deny that we have any genuine cognition at all of ourselves merely through inner sense. Objective cognition of ourselves would have to depend on relating to ourselves as objects by relating to our spatial existence – in other words, to our embodiment.

Kraus decisively rejects the thought that self-cognition must depend on a persisting personal identity dependent on the body. She worries that 'a malfunctioning body-consciousness or just its temporary failure would severely impede my ability to have a consciousness of my personal identity at all. Not being conscious of my own body might also make it impossible for me to experience and later reflect upon a series of dream states' (p. 159).

Well, I should think that malfunctioning consciousness of my body *could* interrupt my consciousness of my personal identity. Others would know I'm the same person during these periods of malfunctioning, but I might not. So Kraus's worries here seem to me unpersuasive. As for dreams: I don't know about you, but I have had dreams where I (the imagined subject of the dream) was not my own waking self but someone else. Even when I do dream that I, Allen Wood, am doing or experiencing various things, I am sceptical that the subject of my dreams is really identical with my waking self. Also, I am sympathetic to the view, shared by Jean-Paul Sartre and by my longtime colleague Norman Malcolm, that dreams are not any kind of conscious experience at all. After all, when I am sound asleep, I am *unconscious.* Dreams are a kind of imaginary object, a kind of fiction. The author of a fiction imagines stories in which the first-person narrator is not identical to the author. Ishmael, for example, is clearly not identical to Herman Melville, and Edgar Allan Poe is clearly not identical to Montresor in *The Casque of Amontillado*, or to the first-person narrators of *The Black Cat* or *The Pit and the Pendulum.* The subject of a dream is bound to a given personal identity even less than the narrator of a fiction. In some dreams who I am even

https://doi.org/10.1017/S136941542200019X Published online by Cambridge University Press

changes in the course of the dream. The only person I can be sure is identical with me is the person who remembers and narrates the dream later on, after I wake up. And that is always my waking embodied self.

How closely did Kant relate self-cognition to our embodiment? Kraus may be correct that the preponderance of textual evidence from his published writings commits him to what she calls a 'psychological account of personhood' (p. 160). Kant clearly wants to remain agnostic about the positions he calls 'spiritualism' and 'physicalism' when it comes to the mind–body relation. He wants to leave open the possibility of both Cartesian dualism and La Mettrie's materialism but also to put both beyond the reach of any philosophical proof. It is unclear, however, how far he can succeed at this, given his own philosophical insights and commitments. Kant admits he is troubled by what he himself calls 'the paradox of inner sense' – that inner sense is not sufficient for persisting personal identity because determination in time must depend on objects of *outer* sense. Yet we seem acquainted with our own selfhood through time entirely from within, and I don't seem to need the testimony of others, or even a look at my body in the mirror or down my shirt-front, to tell me I am the same person I was ten minutes ago. But in treating of the imputability of actions and their consequences, Kant does say that 'I consider myself from the standpoint of another' (A362–3). Kant's B Refutation of Idealism argues that even the determination of time in inner sense requires reference to material objects distinct from my inner perceptions. My own body would seem to be the most readily available and even indispensable object of this kind.

In his published writings, Kant does not draw the explicit conclusion that we can cognize ourselves only by cognizing our body in space. But in some *Nachlass* versions of the B Refutation of Idealism, he does draw this stronger conclusion. The basic argument there is that, in order to determine my states of inner sense in time, I must refer them to objects of outer sense in space. But then Kant continues this line of reasoning as follows: 'We are first objects of outer sense to ourselves, for otherwise we could not perceive our place in the world and thus intuit ourselves in relation to other things . . . I am myself an object of outer intuition, without which I could not know my position in the world. The soul as an object of inner sense cannot perceive its place within the body, but it is in the place where the human being (*Mensch*) is' (*Reflexion* 6315, 18: 619). The human being therefore must be in space, and therefore embodied. Only as embodied can anyone truly cognize him- or herself. Perhaps Kant shrank from these conclusions, or at least wanted to shrink from them; but they do seem to follow from his own arguments.

What is most admirable in Kraus's Kantian reflections is the project of conceiving the topic of the mind in a way that preserves human subjectivity and avoids reductionism in the same way Kant himself does. Kant's approach to the mind is pragmatic in the sense that he understands human beings as active and their self-understanding as part of their agency. The self is an active agent, capable of self-knowledge but a knowledge that is limited by the ways human beings remain self-opaque through the very vocation by which they continually make themselves what they are.

Cite this article: Wood, A. (2022). Comments on Katharina T. Kraus, *Kant on Self-Knowledge and Self-Formation: The Nature of Inner Experience*. *Kantian Review* **27**, 469–474. https://doi.org/10.1017/S136941542200019X