

## AUTHOR MEETS CRITICS

# Subject, Soul and Person in Kant: Questions for Katharina Kraus

Clinton Tolley

University of California, San Diego, CA, USA  
Email: [ctolley@ucsd.edu](mailto:ctolley@ucsd.edu)

### Abstract

Kraus's book is a rich and systematic examination of Kant's account of the different dimensions of the metaphysics, epistemology and phenomenology of the 'self' that pertains to human subjectivity. Here I explore some of the different meanings that Kraus associates with the term 'self' on Kant's behalf, asking for further clarification as to her interpretation of the terms 'subject' ('the I'), 'soul' and 'person', in particular. I also raise some critical questions concerning Kraus's account of the nature and limitations of the 'real' use of the concept of the soul in particular, in light of passages throughout the Critical period in which Kant seems to allow for a relatively unproblematic application of the term to whatever being it is that possesses the various psychical faculties – a being which he also seems to allow is an object of experience (and hence cognition).

**Keywords:** Kant; subjectivity; apperception; psychology; anthropology; idealism; Paralogisms

### 1. Introduction

Katharina Kraus's *Kant on Self-Knowledge and Self-Formation* (Cambridge: Cambridge University Press, 2020; hereafter *KSS*) provides a careful and systematic investigation, along with a nuanced and innovative interpretation, of several overlapping parts of Kant's philosophy of mind and psychology. It is successful to such a high degree that it is surely the 'state of the art' for all future discussion of its topics.

My focus here will be to invite Kraus to clarify her views on the meaning or meanings of 'self' – not least as it shows up in her title – by inviting her to say more about her views on three other concepts that are closely related to that of <self> – namely, the concepts of <the I [*Ich*]>, <soul [*Seele*]> and <person> (I will use angle-brackets to express concepts). One of the many virtues of Kraus's book is that it challenges a common, if often implicit, interpretative assumption that more or less runs these terms together. This kind of interpretation takes its main point of departure from Kant's discussion of 'I' in the Transcendental Deduction, and the fundamental role that 'identity' in consciousness plays in its argument, especially in the B-edition, and sees Kant's argument in the Deduction as requiring appeal to consciousness of 'personal' identity across different representations. Kraus challenges this assumption by foregrounding lessons from the Paralogisms as a criterion for the interpretation of the Deduction itself, which leads her to argue that, already in the Deduction, Kant is

working with a distinction not just between <the I> and the rational concept ('idea') of the <soul>, but also between <the I> and <person> as well, such that the appeals to 'the I' and 'identity' cannot be taken as grounds for seeing anything specifically person-involving at work in the Deduction itself.

Kraus's general arguments for the need for more care with distinctions are quite convincing; here I want to raise more specific questions about how she exactly means to partition the meanings of the relevant terms. First, even if we are sympathetic to Kraus's general case (given in part II especially) for distinguishing between <the I> as picking out something formal or logical, and <person> (and <soul>) as aiming to pick out something substantial, and even if we follow Kraus in noting the close connection between <person> and <soul> that Kant famously draws in the Paralogisms, it is not clear that the only senses in which Kant uses the terms 'person' and especially 'soul' are those which are at issue there. For there would seem to be a 'thinner' sense of 'soul' at work throughout Kant's texts – and in the tradition preceding him – which does not yet imply <person>, and in fact does not seem to be a concept belonging to 'rational psychology' at all. This is the use of 'soul' to refer simply to any kind of substance which has psychical faculties, whether they have merely the lower ones (sensibility, as in 'mere' animals) or also the higher ones (the understanding and reason, as in humans).

Second, and relatedly, even if we are sympathetic to Kraus's very suggestive analysis (given in part III especially) of Kant's ultimate account of the proper use of <person> (and <soul> in the specific sense of an 'idea' of reason) as something which is only 'regulative' of our reasoning about our ordinary self-perceptions and self-experiences, rather than straightforwardly 'constitutive' of any cognition of ourselves as an object, there is still the question of what other term we should use to refer to our own being in the meantime. What should we call this 'thinking thing' – something more than a mere 'form' of some representation, but something which, though not yet a 'person', is nevertheless in possession of faculties sufficient to place them under the 'imperative' to use the concept <person> (and <soul> in its 'rational' sense) to make possible its own 'self-knowledge and self-formation'?

I will proceed as follows. In section 2 I will sketch Kraus's challenges to the common assumption that appeals to personality (via consciousness of personal identity) are at work in the Deduction's talk of 'the I' and 'identity'. I then turn to Kraus's alternate 'formal' interpretation of <the I> in the Deduction as contrastive with the Paralogisms' specifically 'rational' concept of <the soul>, and as implicitly contrastive with <person> as well. In section 3 I turn to the possibility that there is a thinner notion of <soul>, one which is at once more traditional, and also in use and seemingly non-problematically by Kant himself throughout the Critical period. Here I ask whether some such concept of a slightly more substantive object ('thinking thing') might be necessary (or at least useful) for Kraus herself to rely upon, in order to characterize (constitutively) the 'pre-personal' being which is on the road toward self-knowledge and self-formation qua person.<sup>1</sup>

## 2. Subject versus person in the Deduction

Many well-known approaches to Kant's first *Critique* seem to assume that, at least by the time of the Deduction's use of 'the I', Kant means to be using it as a personal

pronoun to refer to some self-same thing that can do and know certain things, and be conscious of itself and its activity, and moreover, that exploiting the awareness of personal identity in particular is crucial to his argument in the Deduction itself. More specifically, many readings have Kant's Deduction depend essentially on the (alleged) consciousness of numerical identity of that being or thing which uses or has the understanding, and so which does the thinking across cases of consciousness or self-consciousness – whether this takes the form of the numerically identical bearer or possessor of a set of representations, or the numerically identical target of the ascription of a set of representations (as all 'belonging to me' or 'in' me; cf. B134; see KSS 92–105).<sup>2</sup>

Kraus's book, by contrast, highlights two kinds of difficulties for this kind of assumption about any alleged person-referring role of 'the I' in the Deduction (in Kraus's words, that 'the I' is 'about *something*'; KSS 96). First, as Kraus notes, from the outset Kant's discussion is almost exclusively conducted in terms of various 'faculties' (*Vermögen*), such as 'the senses', 'imagination', 'understanding' and eventually 'reason' – asking what *they* are, what *they* can do, etc. The main answer to what they can do, in turn, is given in terms of 'representation' (*Vorstellung*). Kant's own exposition therefore starts not with assumptions about the thing that can represent, but rather about the nature of representation itself, the faculties and activities that are involved in bringing it about, in its various kinds.

As Kraus herself clarifies (KSS 19–20), representation is in general understood as that which furnishes the contents of consciousness, and is internally defined as comprising two key constituent relations: a relation to a 'subject' (a subject-directness) and a relation to an 'object' (object-directedness). Kraus calls the relation of a representation to a subject its 'reflexivity', and the relation of a representation to an object its 'referentiality' (KSS 19–21). Though talk of a 'subject' might seem to signal a reference to a person, Kraus emphasizes that the idea of 'subject' in play here is quite minimal and only internally defined relative to any individual representation. For example, nothing yet has been said about whether one and the same ('numerically identical') subject will be present across different representations, or whether the subject has any existence 'in itself', outside of its role or standing in a representation which includes it. The concepts of these further possibilities – i.e. <numerically identical subject across distinct representations> and <subject in itself> – require synthetic additions to the more elementary concept of <subject> which contributes at the ground level to the analytical framework of representation.

The importance of the distinction concerning which concepts do and do not imply a numerical identity comes to the fore in Kraus's recapitulation (in chapter 4) of Kant's discussion of <person> (and <personality>) in the Paralogisms. Here <personality> functions as one of several predicates (along with <simplicity>, <independence ('distinction') from matter>, etc.) which are combined to yield the 'rational' concept ('idea') of 'the soul' of the human being, as it is understood in traditional rational psychology. Like other rational concepts, <soul> in this sense is a concept of an object that lies beyond all experience; its content is distinguished from other ideas by including <immateriality>, <simplicity>, <immortality>, as well as <personality> (cf. B403). Personality itself is here glossed as the property of being an 'intellectual substance'; in the A-edition Kant further specifies the condition of being 'conscious of

the numerical identity of its self (*seiner Selbst*) in different times' (A361; see also 28: 296; cf. KSS 160). <Person> thus goes beyond mere <subject> in implying not just substantiality, but also intelligence (i.e. possession of understanding), as well as consciousness of numerical identity across difference.<sup>3</sup>

Though Kant does mean to be talking about a being in possession of understanding, to the extent that he speaks of what kind of being these faculties are faculties 'of' (i.e. what possesses them), the term 'person' itself does not show up until relatively late – in the B-edition, not until after 400 pages have passed (B403). And if we look to where the concept <person> (and <personality>) finally does make its official appearance, it seems to be identified almost immediately as a 'problematic' concept, since it is one of the predicates taken to characterize the rational concept (*idea*) of <the soul> – as part of what is 'falsely held to be a science of pure reason about the nature of our thinking being' (B403).

Kant takes the primary 'ground' for reason's inference to the personality of our thinking being to be supplied by a representation – namely, 'the representation I' (B404) – which (given the common function of the term as a first-person pronoun) might sound as if it itself already and necessarily characterized its object as a person. Even so, what Kant claims is actually presented via this representation is merely 'the transcendental subject of thoughts = x', a characterization which Kant takes to be neutral between personality and impersonality: 'this I, or he, or it (the thing) which thinks' (B404).

The key insight that allows us to bypass this assumption in the Deduction – one which Kraus both identifies and further clarifies (using resources of contemporary philosophy of language) – is that Kant is quite emphatic about using 'I' only to 'express' the 'mere *form*' present in each case of consciousness, as she puts it later in the work (KSS 106), rather than using 'I' to refer to any *thing* or being beyond what is 'in' consciousness – let alone a full *person*, understood as a being which is conscious of itself as numerically identical across representations or thinkings, or across distinct moments of time. By being only 'expressive' of an element in the form of consciousness itself, the 'I' tracks only a constituent or aspect of conscious representation, rather than anything concerning the being, thing, substance or person which has or possesses this representation. On Kraus's reading, the 'I' of the Deduction expresses solely the awareness of the relation to some subject or other which is present in every conscious representation.

Kraus thus acknowledges (as one must) that Kant does refer to a kind of 'identity' in relation to this 'I' (cf. B133–6). As she convincingly demonstrates, however, Kant does so only to point to the identity (sameness) of *form* across distinct cases of consciousness (conscious representations). Crucially, this is a *qualitative* identity, or uniformity, present in each case of consciousness, on par with the sameness involved, for example, in two distinct judgements being of the same subject-predicate form (<this cat is grey>, <this dog is brown>). Referring to this 'logical' sameness is emphatically not meant to appeal to some *real* or *numerical* identity across (or underlying) cases of consciousness (e.g. that two distinct judgements were made by the same thinking thing or being), despite what many prominent interpretations of the Deduction would seem to assume (cf. KSS 92–105).

### 3. The 'real' use of <the soul>

If the foregoing helps to draw out the textual and systematic motivations behind Kraus's interpretative distinction between merely 'logical' uses of 'the I' and those that intend to refer to a 'real' thing (person), as well as the case for reading the Deduction (in light of the Paralogisms) as depending only on appeals to what is contained in the former, let me now take up her account of the sense in which Kant allows, nevertheless, that there is a legitimate 'real' use of <the I> – or at least, <subject> – to refer to some thing or other. Kraus's own analysis (especially in parts I and II) focuses on Kant's criticisms of the use of this term in the Paralogisms in the context of rational psychology, where 'soul' expresses a concept which is meant to imply <personality> (cf. B700). Even so, in the concluding part III, Kraus attempts to chart out a place for a positive role for <person> to play, in the 'regulation' of our reasoning about our inner experience, as an ideal, and ultimately in an 'imperative' to form our inner experience around such a centering ideal.

One might wonder, however, whether there are other, more 'constitutive' roles for at least some versions of the concepts of <soul> and <person> to play. Concerning 'soul': as alluded to in the first paragraphs above, there are places in the *Critique*, including in the first edition of the Deduction (cf. A94, 124), where Kant uses the term 'soul' to express a concept with a seemingly much less weighty content – roughly, to refer to whatever being it is that possesses the faculties of sensibility, understanding and reason. In fact, these faculties are described throughout the *Critique* as belonging, not necessarily to a person *per se*, but to 'the soul' or 'the mind' (cf. B67–68, 103, 169, etc.). What is more, in such contexts, he seems to use 'soul' both without signalling any hesitation about the problematic status of its object and with the (at least *prima facie*) implication that we certainly know or cognize (at least in some sense) that there is such an object.<sup>4</sup> This use of <soul> would seem to go beyond that of the merely 'formal' sense of <subject> that Kraus herself brings into focus, since the latter is instanced within a given representation and can be used to refer to a merely 'formal' feature of the contents of consciousness, whereas 'soul' here refers to a being or thing which has existence outside of a representation or content of consciousness, as that which possesses specific faculties. Even so, its content would be fairly minimal insofar as it would represent this thing without determining whether it has any of the further rational-psychological determinations (listed above) that reason 'infers' that it must possess. In this respect, Kant's use of 'soul' seems to parallel his use in the Paralogisms of the more neutral 'thinking being (*Wesen*)' (cf. B418) or 'thinking nature' (B418n., 874) – as that being or thing which has the faculty for thinking.<sup>5</sup>

In such a use, Kant might seem to be following a long-standing tradition, still active in early modern German philosophy, which accepts that there is a concept <soul> that is much broader than the 'rational' significance seemingly accorded to the term 'soul' in the Paralogisms. This tradition takes its bearings from Aristotle's own treatment of 'soul' in *De Anima*, which takes the term *psuche* (*anima*) to range over not only non-human animals but also other living things such as plants, which possess merely 'vegetative' soul, despite not being persons.<sup>6</sup> We can see continuing echoes of this in Leibniz's *Monadology* (cf. §§18–30), where 'soul' (*âme*) is used to refer not just to the specifically 'rational soul' distinctive of humans and the divine, but to that possessed by other non-rational animals ('brutes').

Given its contrastingly thinner content, it is perhaps less surprising that there also seems to be a contrast in Kant's assessment of the validity of the application of <soul>, so understood. Consider, for example, Kant's claims in the *Prolegomena*, not just that 'the persistence of the soul during life can be *inferred*', but that 'the persistence of the soul can be *shown* (*dargethan*) during the life of a human being' and, moreover, that this '*proof*' (*Beweis*) will of course be 'granted' (4: 335; my emphasis). The basis for these claims seems to be that, during life, the soul itself is properly thought of as 'an object of *experience*' (*ibid.*; my emphasis), i.e. incorporating what Kant calls elsewhere 'inner experience' (*Anth*, 7: 133–4, 141–2). The sense of 'the soul' that is at issue here obviously cannot be that of the rational concept from the Paralogisms, insofar as this rational concept ('idea') is by definition such that its object cannot be a possible object of experience (compare again B370, 383–4).

What, then, should we take to go into this thinner sense of 'soul'? Beyond the basic affirmation of 'the soul' in this sense as an object of experience, this passage (and others like it) is striking in its further affirmation that *persistence* in particular can be 'proven' of 'the soul', insofar as this would seem to imply (via the First Analogy) that the pure concept (category) of *substantiality* is also valid of 'the soul', again when this term is taken in this experience-immanent sense.

For her part, Kraus shows sympathy instead with those who argue that Kant rejects the ('constitutive') validity of substantiality and other dynamical categories in relation to inner experience, and so also rejects the 'parity' of inner and outer experience with respect to the categories and principles of the understanding (*KSS* 5, 131, 153ff.). Nevertheless, without some such concept – of an object more substantial or thing-like than a mere 'form' of consciousness, even if less determinate than a person – it is not clear on Kraus's account, first, how *else* we should refer to the object of our inner experience, the being which thinks and infers and regulates its activities so as to approach that 'self-knowledge' and 'self-formation' which would lead to becoming a person.<sup>7</sup> It is also not clear why Kant would not be more scrupulous in his use of 'soul' and more emphatic in his rejection of its applicability to inner (intuition, sense and) experience.

Finally, concerning 'person' itself: if we add to the foregoing the fairly minimal conception of 'person' given in the A-Paralogisms – roughly: 'substance with intellect' – and if we allow that Kant takes us to be able to know that, whatever else we are, we are in fact souls (in this thin sense) which have intellects, then it is hard to see how we could not also know that we are 'persons', again in this thinner sense. Kant can seem to claim this more or less explicitly in the Antinomy, affirming that we 'cognize' the faculties of understanding and reason in ourselves (B574–5; cf. *G*, 4: 452).

None of this speaks directly against Kraus's criticisms of attempts to see Kant already within the Deduction as appealing to this kind of knowledge or cognition of our personality (i.e. numerical identity), rather than to formal features (qualitative and logical identity) in representations. Nor need it stand as an obstacle to the basic picture sketched in the final chapters of Kraus's book, in which she develops the innovative position according to which Kant means for the specifically 'rational' concepts of <soul> and <person> to function not as representations of properties that we can know already obtain in the constitution of any existing object, but instead as 'guiding threads' for the regulation of certain activities which transform a certain kind of being into a person. It does raise questions, however, both about whether Kant is

more committal about the ontological make-up of the kind of being that is on the road to personhood, and about whether there isn't a distinct (and quite traditional) use of 'soul' (and thinner sense of 'person') which would be appropriate for referring to just this kind of being – one which possesses faculties for thinking (and willing) and which has the potential for becoming a person, and perhaps even (ideally) an immortal soul.

## Notes

- 1 At the outset Kraus uses 'person' to refer to both (a) that which is on the way to being a person (but is not yet), and also to (b) that toward which such a being is regulatively forming itself (cf. KSS 1–2), though she concedes at the end that this would seem to involve her in a kind of 'circularity' (KSS 269).
- 2 Kraus herself takes up Kitcher, Keller, Longuenesse and Wuerth as holding views in this neighborhood.
- 3 Kraus herself points to three related properties requisite for being a person: being a subject (i) that itself possesses *representational faculties*, and (ii) has a *perspective* on its representations, and (iii) has a power of *self-determination* (KSS 10). Kant's reference here to 'intelligence' seems more restrictive than possession of some 'representational faculty' or other.
- 4 Compare also Kant's occasional description of 'inner intuition' as intuition 'of our souls' (B50); cf. KSS 8, n. 19.
- 5 Kant seems to acknowledge the alignment of the two phrases ('thinking being' and 'soul') at the outset of the Paralogisms: 'I, as thinking, am an object of inner sense, and am called "soul"' (B400).
- 6 Kraus herself alludes to this broader concept toward the end of her book (KSS 260), noting the older Aristotelian distinction between <soul> as such (qua *anima*) and the more specific <human soul> (qua *animus*, or 'rational animal').
- 7 By rejecting the validity of 'the soul' in even this thinner sense, Kraus would seem to make not just personality but even substantiality (a category of the understanding) and persistence into merely 'regulative' (rather than constitutive) concepts for inner experience; cf. KSS 155–6, 202–3. By contrast, accepting the use of 'the soul' in this thinner sense allows us to concede empirical cognition of 'the soul', and restrict the regulation of inner experience solely to the properly 'rational' concept of <soul> (qua 'idea').