The Third Paralogism and the Transcendental Unity of Apperception

This week we are focusing only on the 3rd of Kant’s Paralogisms. Despite the fact that this Paralogism is probably the shortest of the four and in some respects its argument can summarized more quickly than some of the others its implications are more far-ranging than is initially obvious due to the fact that understanding the argument requires us to revisit one of the central aspects of the Transcendental Deduction, namely, the status of the transcendental unity of apperception.

The 3rd Paralogism treats a claim concerning personality though personality is presented here as connected to consciousness of numerical identity at different times. That which is conscious of numerical identity at different times is said to be a person and, it is claimed, that since the mind or soul is so conscious of itself that what has a mind or soul is therefore a person. In responding to the claim of the argument of those advancing the position of the 3rd Paralogism Kant begins by discussing the way in which phenomenally I relate to the numerical identity of an object that is external to me, which is in terms of a permanent element in it that every determination of it is related to. This permanent element is what is taken to be identical throughout the changes of the object.

Having begun with this preliminary observation Kant moves to on to discussing the case of the self. In this case awareness is of something that is inner and there is a condition under which any item of inner sense can be
perceived, which is the form of time. So in order to relate my states to a
sense of myself I do so through succession. This leads Kant to his first
substantive point in response to the argument of the Third Paralogism which
is to point out that the personality of the soul or mind is not something that
is inferred from the data of inner sense as one would suppose from the
argument that is given. It is not so inferred as to be aware of the self is
immediately and analytically to be aware of something identical over time.
To be aware of self-consciousness in time, he states, is to be aware of
something that is strictly identical, an awareness that is a priori but only due
to its also being analytic.

Kant’s first response to the argument of the 3rd Paralogism is thus to
deny that the argument is needed in the sense of attempting to provide an
inference for something that is merely analytically given. This point is
amplified when Kant goes on to state that identity is unfailingly met with in
my consciousness of myself. Whilst this provides a reply of sorts it would
be unlikely to convince anyone who held the view stated in the 3rd
Paralogism. So Kant next changes tack moving from my sense of my
identity to the view of an external observer of me. This external observer
could admit that the representation of an “I” always accompanies my
consciousness without thereby being led to the conclusion that there is a real
permanence in me. This contrast between the perspective of the observer
and my own shows that: “the identity of the consciousness of myself at
different times is therefore only a formal condition of my thoughts and their
coherence, and in no way proves the numerical identity of my subject” (A363). In other terms, Kant is admitting here much what he did when treating the 1st Paralogism, which is that whilst the “I” is the subject of all thoughts, this is not sufficient for it to be a substance (something permanently identical).

The status of the claim to say “I” over time involves a logical identity of meaning but underneath this identity of meaning there may be concealed a difference of reference. In order to illustrate this point clearly Kant presents a striking argument from analogy that is set out in a footnote to A363-4. Here he invites us to conceive of an elastic ball that hits another similar ball and in the process conveys its whole motion to the second ball. Just as this is conceivable, so he suggests, it could also be the case that one mental substance could communicate all its states to another with this second doing so to a third in such a way that all memories and consciousnesses are being transferred. Thus the last such substance would have a sense that it had been aware of all the previous states and memories as its own even though there would have been no constant identical thing given.

Our consciousness of our own identity thus gives no substantive guarantee of permanence. We cannot establish that the “I” which unites together diverse representations over time has therefore an equivalent referential range as the thought of the “I” may well be as much in flux as any other thought we can have. Permanence of selfhood does not have any
basis for being attributed to us as something that produces the
transcendental unity of apperception or is its foundation. Rather it is only
due to the transcendental unity of apperception that we arrive at a statement
of permanence of the self. So what we discover in the 3\textsuperscript{rd} Paralogism is that
statements of identity of consciousness over time are insufficient to lead us
to attribute a permanence to the bearer of such the states described by these
statements.

Rather than see personality as a determination that attaches to
something of which we are aware therefore we can attach it, at best, to the
transcendental ground of that which could correlate to the unity of the
subject of which we are aware but which we can know nothing of. Viewed
in such a transcendental way it may well have a role to play in practical
philosophy but not as some claim to self-knowledge by means of pure
reason.

Having made clear the basic argument of the 3\textsuperscript{rd} Paralogism I now
want to develop some reasons why this argument has caused some
perplexity amongst readers of the \textit{Critique}. This concerns the suggestion,
based on the discussion of the transcendental unity of apperception in the
Transcendental Deduction, that Kant in fact requires a view of personal
identity that is remarkably akin to the one he is apparently rejecting in his
critical response to the 3\textsuperscript{rd} Paralogism. The reason for holding such a view is
not difficult to find. For example at A108 Kant speaks of an “original and
necessary consciousness” of “the identity of the self”, an identity that he
goes on to say we can think \textit{a priori} as productive of a transcendental unity that is the ground for \textit{a priori} rules that are at work in making our relation to appearances law-governed. So it is natural to take this as an argument for a form of personal identity, a form of personal identity that appears to have the structural role in the theory of the A-Deduction that is specifically being criticized in the 3\textsuperscript{rd} Paralogism, particularly when we add that A107 Kant speaks of the transcendental unity of apperception as necessarily having to be represented as numerically identical.

There are, however, two ways of taking these statements. Rather than view them in the way I have just suggested it is possible to see them instead as indicating that what is identical is something formal and that it is the unity of this formal apperception that has be to seen as identical. This would fit passages in the B-Deduction that seem to parallel those just cited from the A-Deduction as when at B134 Kant states: “only in so far as I can grasp the manifold of the representations in one consciousness, do I call them one and all \textit{mine}” (B134). Shortly after making this statement Kant also goes on to say that the principle of the necessary unity of apperception is an identical or analytic one, a point which connects with the manner of his response to the 3\textsuperscript{rd} Paralogism. On the basis of these considerations Karl Ameriks thus argues that Kant is denying that there is a relationship of mutual entailment between personal identity and transcendental apperception. I will now turn to some of the grounds for accepting this point.
Firstly, in the A-edition treatment, Kant, after going through all four of the paralogisms, turns to the question of how the illusions that are at work in them confuse us. Here he first discusses the question of what the constitution is of a thing that thinks (A398) indicating that we can give no *a priori* reply to the question. The reason why this is so is subsequently given: “In the absence of an underlying intuition the category cannot by itself yield a concept of an object; for by intuition alone is the object given, which thereupon is thought in accordance with a category” (A399). But there is no such intuition given of transcendental apperception and hence no sense of personal identity as objective can emerge from it. The reason why Kant claims this becomes clear subsequently when he relates transcendental apperception to the categories stating that the “I” does *not* know itself by means of the categories but rather: “knows the categories, and through them all objects, in the absolute unity of apperception, and so through itself” (A402).

It is by means of relation of representations to the transcendental unity of apperception that we form a concept of objects so if the transcendental unity of apperception is a condition of the representation of objects then it cannot itself be known as an object. “Nevertheless there is nothing more natural and more misleading than the illusion which leads us to regard the unity in the synthesis of thoughts as a perceived unity in the subject of these thoughts.” (A402). It is this illusion that produces the 3rd Paralogism. This point is repeated in the B-edition treatment where the
illusion underlying the 3\textsuperscript{rd} Paralogism is identified with that underlying the 1\textsuperscript{st} (B421-2) and Kant here adds: “The subject of the categories cannot by thinking the categories acquire a concept of itself as an object of the categories” (B422).

This claims come in the conclusion of Kant’s treatment of the whole of the Paralogisms although there is a similar point made when Kant begins his discussion, where he refers to the simple representation of the “I” at A346/B404 saying of it is not even a concept but only that “a bare consciousness which accompanies all concepts” and adding that: “It is known only through the thoughts which are its predicates, and of it, apart from them, we cannot have any concepts whatsoever”.

Despite making these points as the indication of the basis of the illusion that is at the root of the 3\textsuperscript{rd} Paralogism however Kant also elsewhere in the \textit{Critique} reinforces the point that there is a practical sense that can be given to personality in terms that suggest that perhaps the mistaking of practical claims for theoretical ones might be the root of the problem. Thus, for example, at A546-7/B574-5 Kant contrasts two ways we have of “knowing” ourselves writing: “Man…who knows all the rest of nature solely through the senses, knows himself also through pure apperception; and this, indeed, in acts and inner determinations which he cannot regard as impressions of the senses. He is thus to himself, on the one hand phenomenon, and on the other hand, in respect of certain faculties the action of which cannot be ascribed to the receptivity of sensibility, a purely
intelligible object.” This latter notion is however later explained in terms of our ability to apply moral imperatives to our conduct so that Kant here moves from the sense of pure apperception in a theoretical sense to a notion of autonomy in a practical respect. Unless some slippage of this sort was to happen in any case the reference to “knowing” the self by means of transcendental apperception would tilt directly against the distinction between the self of appearance and the self-in-itself set out in the B-Deduction (B157-9).

So perhaps the illusion underlying the 3rd Paralogism is double in form and even difficult to avoid for Kant. What seems finally to follow from Kant’s treatment of the 3rd Paralogism is that there are no transcendental criteria of a theoretical sort that could ever securely identify the self and, since the empirical self is subject to the difficulties that Hume analysed, it would seem further that this entails that there is fundamentally only a practical sense can be given to personal identity, a sense that requires therefore that there be an alternative form of metaphysics to that which can be given in theoretical philosophy, a metaphysics of morals.