On Sellars’ Exam Question Trilemma: Are Kant’s Premises Analytic, or Synthetic A Priori, or A Posteriori?

James R. O’Shea, University College Dublin

(This is the author’s post-peer reviewed publication version. For citation please refer to the published version (2019) in the British Journal for the History of Philosophy, 27:2, 402-421. DOI: 10.1080/09608788.2018.1524368.)

* * *

Wilfrid Sellars is now widely recognized to be one of the most important of those twentieth-century analytic philosophers who argued that Kant’s transcendental philosophy contains deep insights that can be adapted to address our own philosophical problems. Sellars expressed this outlook when he wrote that “Kant’s account of the conceptual structures involved in experience can be given a linguistic turn” and thereby “be seen to add essential elements to an analytic account of the resources a language must have to be the bearer of empirical meaning, empirical truth, and . . . empirical knowledge” (Sellars 2002a OAPK §31). In what follows I propose to examine the methodological aspects of Kant’s transcendental account of our “conceptual structures” that, I contend, Sellars took to be fundamental not only to Kant’s view but to influential themes in his own philosophy. The topics that I will focus on throughout, however, are not primarily those pertaining to Sellars’ more famous (or infamous) comparisons of Kant’s appearances/things in themselves distinction to Sellars’ own manifest image/scientific image distinction (cf. Sellars 1968; O’Shea 2018). My first aim here is to clarify in some detail and to argue for the plausibility of Sellars’ particular interpretation of what he takes to be Kant’s “analytic” transcendental method in the Critique of Pure Reason (“analytic” in the sense that Kant’s own method is interpreted by Sellars as based ultimately on claims that are supposed to be analytic a priori conceptual truths), in particular pursuing some of Sellars’ references to the “Doctrine of Method” at the end of the first Critique. This requires an examination of the nature of Kant’s “transcendental proofs” of synthetic a priori (SAP) principles, and in particular of how such proofs avoid a certain methodological trilemma of (1) an SAP regress (or question-begging circularity),

1 I address the general topic of Kant’s relationship to twentieth-century analytic philosophy in general in O’Shea 2006, and then with particular focus on Sellars in O’Shea 2016 and forthcoming; the present quotation and a couple of short passages will also appear in the latter publication).

2 I will follow the standard practice of referring to Sellars’ works by means of their abbreviations and with section or paragraph numbers where possible. In this case “OAPK” refers to Sellars’ “Ontology, the A Priori and Kant,” collected in Sellars 2002a KTM by Jeffrey Sicha. Sicha explains (p. 261) that OAPK was an unpublished typescript from the mid-1960’s (apparently revised by Sellars in 1970) which was to be ‘Part One’ of a longer piece, ‘Part Two’ of which was published as “Kant’s Theory of Experience” in 1967 (KTE, also in Sellars 2002a).

3 Sellars uses our linguistic behavior as his model for understanding the nature of inner conceptual thinking, whereas Kant’s focus is directly on the logical form of concepts and judgments themselves. However, given that in all relevant respects what Sellars wants to say about language he also wants to say, by analogy, about inner conceptual thinking, I will assume, at least for the purposes of this paper, that no obstacles arise from Kant’s primarily non-linguistic methodological approach to the nature and role of concepts in our cognition.
(2) merely analytic \textit{a priori} triviality, or (3) a self-defeating reliance upon \textit{a posteriori} hypotheses.\(^4\) I then examine in part II how Kant’s method, as interpreted in the analytic manner explained in part I,\(^5\) was adapted by Sellars to produce some of the more influential aspects of his philosophy, expressed in terms of what he contends is their sustainable reformulation in light of the so-called linguistic turn.

\section*{I. Sellars’ “Analytic” Interpretation of Kant’s Transcendental Method}

In the introduction to the first \textit{Critique} Kant tells us that he calls “all cognition \textit{transcendental} that is occupied not so much with objects but rather with our \textit{a priori} concepts of objects in general,” and that a “system of such concepts would be called transcendental philosophy” (A11–12/B25). For Sellars, Kant’s methodological shift from the direct investigation of reality that was characteristic of traditional metaphysics to the “transcendental” attempt to uncover the \textit{a priori} principles that necessarily structure our sensory apprehension and conceptual understandings of objects, represented a decisive methodological step forward.

Following Kant’s own frequent practice, let us take the \textit{a priori} conceptual category and corresponding principle of \textit{causality} as our lead example of the sort of thing that Kant means by a transcendental cognition in the above sense. By what general method or approach does Kant attempt his “proofs” of such \textit{a priori} transcendental principles as the famous “\textit{Principle of temporal sequence according to the law of causality}” in the “second analogy” of experience, i.e. that “all alterations occur in accordance with the law of the connection of cause and effect” (B232)? Like Hume, Kant holds that evidence derived \textit{a posteriori} from the uniformities of past and present experience fails to account for the objective necessitation and the universality that the causal principle represents as governing the objects of our experience. Similarly Kant agrees that no conceptual \textit{analysis} of the subject-term of a causal relation will logically entail the concept of its

\footnote{What I am calling Sellars’ “analytic reading” of Kant is thus to be understood as referring to Sellars’ embrace of the second horn, mentioned in the main text here, of his own exam question trilemma to be discussed below. That is, I argue in favor of Sellars’ view that the ultimate premises of Kant’s transcendental “proofs” are conceptual analyses in the sense of consisting of non-trivial analytic \textit{a priori} truths or judgments. On Sellars’ reading, as we shall see, such “philosophical” knowledge roughly speaking consists in “analytic knowledge about synthetic knowledge” (involving the complex analysis, in particular, of the \textit{concept} of an \textit{object of possible experience}), while nonetheless also avoiding the other two “synthetic” horns of the trilemma. The resulting ineliminable ambiguity in Kant between relatively trivial analyses of given concepts (e.g., “bodies are extended”) and significantly non-trivial conceptual analyses (such as those characteristic of Kant’s transcendental analytic in general, which Kant also describes as an analysis of the \textit{faculty of understanding} (A65–66/B90–91)), will be noted as the argument proceeds.}

\footnote{Sellars’ analytic reading of Kant’s transcendental method as discussed in this paper does not refer to, though it can be squared with, Kant’s well-known methodological distinction between the “analytic method” he is following in the \textit{Prolegomena} as opposed to the “synthetic method” that he indicates he followed in the first \textit{Critique} (cf. Kant 1783, 4.263–4; 4.274–5). A careful recent analysis of the latter distinction is Gava 2013, building on Bird 2006 and other commentators, which I take to be consistent with Sellars’ interpretation of Kant’s method in the first \textit{Critique}. Along one dimension of that distinction, Kant in the \textit{Prolegomena} assumes more directly the validity of the SAP principles to be found in mathematics and natural science than he does in the \textit{a priori} synthesis-revealing analyses of the pure forms of sensibility and understanding provided in the first \textit{Critique}. Sellars’ reading is also consistent with Gava’s further argument that both of Kant’s analytic and synthetic methods are to be found within the first \textit{Critique} as a whole.}
predicated effect, or vice versa. Hence Kant states that the “real problem of pure reason is now contained in the question: How are synthetic judgments \textit{a priori} possible?” (B19).

In both the transcendental deduction (e.g., A92–4/B124–7) and in his account of the principles of pure understanding (e.g., A154–8/B193–7), Kant explains that in the case of such \textit{a priori} synthetic judgments and principles it is the “\textit{possibility of experience}” that “gives all of our cognitions \textit{a priori} objective reality” (A156/B195; cf. A93–4/B125–6, A786/B814). Earlier in the transcendental aesthetic, Kant had of course already argued that all our cognition of “inner” and “outer” objects is a receptive sensory experience of those objects as structured by the pure forms of intuition and hence as necessarily situated in one three-dimensional space and in a single time-order. Kant’s views on sensibility and the pure forms of intuition, space and time – including his signature \textit{transcendental idealist} thesis that we can have knowledge only of appearances or phenomena, not of things in themselves or noumena, not to mention the recently much discussed “conceptualist” and “nonconceptualist” aspects of Kant’s view – were explored in innovative if controversial ways by Sellars.\footnote{See especially Sellars 1968 SM, Chs 1–2. See also deVries 2005, Haag 2017, Landy 2015, McDowell 2009, O’Shea 2016, 2018, Rosenberg 2005, and Westphal 2010. See also, for a substantial investigation of many of the topics I will be exploring in this chapter, Jeffrey Sicha’s book-length introduction to his edition of Sellars 2002a \textit{KTM}. For recent work on Kantian conceptualism/nonconceptualism debates, see especially Schulting, ed. 2016.}

We shall return to the crucial role of sensibility below, but the aspects of the Sellarsian reception of Kant’s philosophical method that I will begin by focusing on concern those schematized principles of understanding that integrate the pure forms of thinking and sensibility in application to experience, and the objective validity of which, Kant argues, makes possible our experience of mind-independent, persisting and causally related material substances within one unitary nature.\footnote{Sellars is among those interpreters of Kant who hold that the transcendental aesthetic cannot fully adequately accomplish its tasks, including in particular the explanation of the nature of our singular intuitions, until the results of the subsequent transcendental analytic are in hand (cf. Rosenberg 2005, 62–63: “The understanding turns out to be implicated in \textit{all cognition}, singular as well as general”). As Sellars puts it informally in his undergraduate Kant lectures, the “trouble is that Kant never went back and re-did everything that should have been re-done to bring the \textit{Aesthetic} into keeping with the requirements of the theory of experience as he develops it in the \textit{Analytic}” (Sellars 2002b \textit{KPT}, 124; though he also remarks informally (KPT 178n): “What the \textit{Aesthetic} does is to establish the noninferential character of our consciousness of spatial objects. What the \textit{Principles} do is to establish that the objects do have the kind of structure which entitles us to call them a physical world.”). For recent contrasting discussions of how to approach Kant on sensible intuition in the \textit{Aesthetic} in relation to the \textit{Deduction}, see Allais 2017 and Conan 2017. For the idea that Sellars represented a complex middle way on the current Kantian “conceptualism vs. non-conceptualism” controversies, differing in important ways, for example, from John McDowell’s conceptualist outlook, see O’Shea \textit{forthcoming}, §3, as well as Landy 2015.

So again: by what method does Kant attempt to prove such synthetic \textit{a priori} transcendental principles? Consider the following challenging examination question that Sellars used to set for the undergraduate students in his Kant class at the University of Pittsburgh in the 1970s, in the form of a trilemma that confronts any candidate premises for Kant’s transcendental proofs (italics added):\footnote{There are of course many closely related discussions by earlier generations of commentators concerning the question of Kant’s starting points and what premises he might be relying on in his transcendental proofs and in the transcendental deduction. To mention just three: Stephan Körner’s 1967 (cf. 1966) would have been well-known to Sellars and raises an interestingly related trilemma objection, to which Körner (unlike Sellars) famously argues that Kant has no good answer (cf. Westphal 2004: 16–17); Kemp Smith’s earlier...
“Either Kant’s premises are proposed as analytic, in which case, if sound, his conclusions must be analytic and hence empty;

or, his premises are proposed as synthetic a priori, in which case he is begging the question;

or, his premises are proposed as synthetic a posteriori, in which case, if sound, his conclusions must be empirical hypotheses.

Discuss.”

Call the three options presented by the trilemma the “analytic” reading, the “synthetic a priori” or SAP reading, and the “empirical” or a posteriori reading of the nature of the starting points or fundamental premises of Kant’s transcendental deductions and proofs. This question raises fundamental and longstanding questions of Kant interpretation, and it is not immediately obvious how Sellars would have addressed his own Kantian methodological trilemma. But I think we can assemble some key aspects of his interpretation, which I suggest furthermore (though not uncontroversially) characterized Sellars’ own distinctively Kantian approach to philosophical problems, expressed in ways that reflected his having developed the “linguistic turn” or the “new way of words,” as he was wont to put it, in the particular ways that he did.

Sellars himself bluntly ruled out the SAP option:

It is obvious to the beginning student that the truths of ‘transcendental logic’ cannot themselves be ‘synthetic a priori’. If they were, then any transcendental demonstration that objects of empirical knowledge conform to synthetic universal principles in the modality of necessity would be question-begging. (KTE §10)

It should be noted, however, that Colin Marshall has recently defended what he presents as a non-question-begging version of the SAP reading such that Kant’s “own explanations take some synthetic a priori claims as explanatorily basic” (2014, 551, italics added), but not the SAP judgments that Kant attempts to prove from those basic resources. Among the examples Marshall examines are the principle of contradiction, which he argues is not analytic for Kant; the fundamental distinction between intuitions and concepts itself; the fact that we have just so many 1923 commentary, which Sellars admired, raises partly related queries concerning the nature of Kant’s starting point (cf. 241–2, for example); and much earlier, the post-Kantian Jakob Friedrich Fries had raised what came to be known (via Karl Popper) as Fries’ or the Friesian Trilemma, and which has recently been examined carefully by Sperber 2017. See also Marshall 2014 for what he calls the “Regress Argument” concerning Kant’s explanation of synthetic a priori judgments, which is laid out in terms of what amount to the same basic choices as Sellars’ exam trilemma, along with helpful citations to earlier literature. In many cases this sort of challenge has led to the conclusion that Kant failed to ground his own philosophy adequately at this fundamental level (cf. Forster 2008, Chs. 11–12). In addition, of course, there is a voluminous literature concerning the nature and fortunes of “transcendental arguments” in general: for a start, cf. Stern ed. 1999 and Ameriks 2003, Chs. 2–3. For my present purposes here I will focus primarily on the question of how Sellars interpreted and appropriated Kant’s transcendental method. (My thanks to Willem deVries for a copy of Sellars’ undergraduate exam questions on Kant.)
logical functions of judgment; and that we have the particular forms of sensible intuition that we do have (2014, §§2, 4). Marshall’s reading attributes to Kant “a short-range rationalist reflection” or “act of inner attention that justifies at least some claims concerning formal features of the mind,” arguing that “Kant accepts a doctrine of limited mental transparency” (2014, 566). Marshall indicates that his reading agrees with Karl Ameriks in holding that Kant was thus “advancing a modest form of rationalism,” though Marshall takes himself to be bringing out a basic rationalist methodological dimension that underpins Kant’s view, whereas “Ameriks thinks Kant’s methodology just involves deducing necessary conditions for ‘common experience’” (566n; Marshall cites Ameriks 2003, 4–12).

While I cannot examine Marshall’s textual evidence and arguments in detail here, we can contrast his approach with what Sellars’ alternative reading of Kant’s method seems to have been and explore the plausibility of the latter. Sellars follows his curt dismissal, above, of the SAP option with the following claim:

It must in a tough sense be an analytic truth that objects of empirical knowledge conform to logically synthetic universal principles. It must, however, also be an illuminating analytic truth, far removed from the trivialities established by the unpacking of ‘body’ into ‘extended substance’ and ‘brother’ into ‘male sibling’. (KTE §10)

This makes it clear that Sellars embraced what in some sense must be a version of the analytic reading or approach to the trilemma, an approach which faces its own textual and philosophical hurdles. In fact, Marshall dismisses the analytic option (which he takes to be exemplified by Strawson 1966) in just two short sentences, appealing to Kant’s clear view that “synthetic judgments . . . can by no means arise solely from the principle of analysis” (Prolegomena, 4:267; and cf. A216–18/B263–5). Marshall also cites Kant’s opening remarks in the Analytic of Concepts, “where Kant contrasts his project of analyzing the faculty of understanding with the more common project of analyzing concepts” (2014, 557n29; see also Land 2017, 510–11). Here is Kant:

I understand by an analytic of concepts not their analysis, or the usual procedure of philosophical investigations, that of analyzing [or “dissecting”: ergliedern] the content of concepts that present themselves and bringing them to distinctness, but rather the much less frequently attempted analysis [dissection] of the faculty of understanding itself, in order to research the possibility of a priori concepts by seeking them only in the understanding as their birthplace and analyzing its pure use in general; for this is the proper business of a transcendental philosophy; the rest is the logical treatment of concepts in philosophy in general. (A65–66/B90–91)

However, we shall see that a non-trivial analytic reading can be seen to rely on, rather than being clearly defeated by the distinction Kant highlights in this passage, i.e., between the “usual procedure” in “philosophy in general” of analyzing given concepts “that present themselves” by “bringing them to distinctness,” as opposed to Kant’s own transcendental investigation of the possibility of a priori concepts by analyzing the “pure use” of the understanding in general.9

---

9 The analyses of the a priori concepts (categories) of the faculty of understanding in its “pure use” as making experience possible are thus to be distinguished, for example, both from the analysis of concepts given a
In the metaphysical deduction or “clue to the discovery of all pure concepts of understanding” that immediately follows the passage quoted above, Kant’s analysis of the pure use of understanding draws out, inter alia, that “concepts rest on functions” understood as “the unity of the action [or act] of ordering different representations under a common one,” and that “there is no other use of these concepts than that of judging by means of them,” concepts thereby being clarified as “predicates of possible judgments” (A68/B94). No doubt it can be argued in accordance with an SAP reading that “the act” of mind described above requires a moderate rationalist synthetic a priori grasp of the mind’s own synthetic “actions,” involving a thesis of mental transparency. However, it can alternatively be argued plausibly that by thus clarifying that concepts are general representations of objects in virtue of their essential role in judgments as “functions of unity among our representations,” so that concepts “as predicates of possible judgments, are related to some representation of a still undetermined object” (A69/B94), Kant is then able to offer in particular a “transcendental” (as opposed to merely “logical”) analysis of how a priori concepts or categories of understanding function in their “pure use” in the possible cognition of an object – an “object” which is then further clarified as an object of possible experience as the transcendental analysis proceeds (cf. §19 of the B-Deduction in particular). In comparison with what Kant views as the starting-point options presented by his empiricist and rationalist predecessors, the above transcendental analysis of concepts as functioning in the understanding’s judgments about objects plausibly meets the test of being “illuminating” or pregnant with further substantive philosophical consequences, while also arguably not resting either on unquestioned rationalist SAP assumptions or on the sorts of analysis of the content of given concepts that Kant regards as posterior to this properly critical, transcendental analysis.

The several notes that Kant wrote in his copy of the first edition of the Critique accompanying the passage cited above (A65/B90) lend further support to the view that by the “analysis of the faculty of understanding itself” Kant means “the analysis [Analysis] of experience contains, first, its analysis [Zergleiderung] insofar as judgments are in it; second, beyond the a posteriori concepts also a priori concepts,” which includes showing that “experience” or “empirical cognition,” which “consists of judgments,” “is only possible through synthetic a priori propositions.” This and the other texts examined above suggest that Kant’s transcendental analysis takes our judged experience, as involving the concept of an object of our possible experience or empirical cognition in general, to be what is being analyzed in such a way as to display that SAP judgments and principles such as the causal principle are necessary for the possibility of such cognition.

Sellars’ analytic reading of Kant’s method accordingly stresses a point that Sellars repeatedly emphasizes, that Kant was “not attempting to prove that there is empirical knowledge, but to articulate its structure” (1967 KTE §45; cf. KTE §11; TTC §§51–6). This invites comparison with Ameriks’ much discussed argument a decade later in “Kant’s Transcendental

posteriori, and also from the traditional analyses of the concepts “that present themselves” in rational metaphysics as allegedly given a priori without the need for any prior investigation of their possibility or of the objective validity of their “pure use” in relation to the objects of experience.

10 These notes were published by Benno Erdmann (1881) and were translated and included in the Guyer and Wood translation of the Critique listed in the references. The Erdmann and Akademie reference here is E XXXIII, pp. 21–2; Ak. 23:24–5. Kant’s overall analysis of the concept of an object of possible experience here begins with an analysis of the nature of judgment and attempts to clarify the functional role of concepts “as predicates of possible judgments” (A69/B94), thereby exposing by means of the analysis the pure a priori forms of judgment and conception that make our a posteriori empirical cognitions of objects possible.

11 See also Sellars 2002a KTM, 483–4 (= Sellars’ “Cassirer Lecture Notes,” CLN B38–41).
Deduction as a Regressive Argument” (in 1978) against what he called the “Received Interpretation” of the goal of Kant’s transcendental deduction. The Received Interpretation, roughly put, is that Kant was attempting to defeat the Humean sceptic through a deductive argument from non-epistemic premises concerning subjective consciousness and our mental representations to the conclusion that we have empirical knowledge of an objective, SAP-lawful world. Ameriks argues to the contrary that “the deduction” should be seen “as moving from the assumption that there is empirical knowledge to a proof of the preconditions of that knowledge,” that is, “to the universal validity of the categories” and of synthetic a priori principles (2003, 51, 66). Sellars agrees that Kant is not attempting to refute radical sceptical arguments directly on their own terms by justifying our beliefs in the face of such challenges (cf. Bird 1974). He stresses, however, that Kant’s analysis of the concept of empirical knowledge “is such as to rule out the possibility that there could be empirical knowledge not implicitly of the form ‘such and such a state of affairs belongs to a coherent system of states of affairs of which my perceptual experiences are a part’,” where by such a “system” Sellars means Kant’s conception of a directly perceived and objectively lawful system of nature, as opposed to the sorts of allegedly incorrigibly known states of consciousness to which Descartes or Hume appealed. Kant’s transcendental analysis thus indirectly undercuts both the skeptic and the ‘problematic idealist’ who, after taking as paradigms of empirical knowledge items that seem to involve no intrinsic commitment to such a larger context, raise the illegitimate question of how one can justifiably move from these items to the larger context to which we believe them to belong. (Sellars KTE §11)

In holding that Kant’s transcendental method is based fundamentally on an analysis of the “complex concept” of empirical knowledge (KTE §9), is Sellars in effect combining the “analytic” and the “synthetic a posteriori” (empirical) responses to the exam question trilemma above? The exam question portrayed the a posteriori horn as entailing that Kant’s “conclusions must be empirical hypotheses,” but this consequence would follow only if the supposed empirical premises consisted in the sorts of empirical facts that either are, or are appropriately explained in terms of, empirical hypotheses. However, the general anti-sceptical starting point of Kant’s analytic method, namely that we have empirical knowledge in general, is warranted not by any empirical or explanatory hypothesis but rather indirectly by the misconceptions that the analysis reveals must plague any attempt to start out from premises that allegedly prescind from any such objective cognitive commitments in experience.

In support of the interpretation sketched above, Sellars remarks that “Kant’s discussion of philosophical method in the concluding chapters of the Critique shows that he was fully aware of these facts, and realized as well that ‘transcendental logic’ as knowledge about knowledge could consist of analytic knowledge about synthetic knowledge” (KTM §10 n2, italics added). Unfortunately, Sellars here provides no specific references to what passages in the Critique’s multifaceted “Doctrine of Method” chapters he has in mind. However, we can find such references in his notes for his 1979 Cassirer Lectures at Yale, where he again asks whether “the propositions of Transcendental Metaphysics [are] analytic or synthetic? (They are a priori.)” He continues: “Kant says that knowledge of principles through concepts only (as opposed to based on experience) is always analytic” (CLN in 2002a, 483–4); he then notes Kant’s assertion in the Discipline that “Philosophical knowledge is the knowledge gained by reason from concepts” (A713/B741),12 in order to draw his

---

12 The full sentence in the Guyer/Wood translation is: “Philosophical cognition is rational cognition from concepts, mathematical cognition that from the construction of concepts.”
conclusion that “therefore” philosophical cognition, for Kant, “must be analytic” (ibid., 484). Of course, this definition, if considered just by itself, would equally apply to the “dogmatic” rationalist analysis of concepts in Leibniz, as Kant himself makes clear in a note from the 1780s on the same topic. But following up Sellars’ two further page references to Kant’s text in this context clarifies the matter: “…Therefore, [philosophical cognition] must be analytic. We can’t know the causal principle through the concepts of event and cause. We need the concept of ‘possible experience’. (See A736–7/B764–5 and A788/B816)” (Sellars 2002a, 484).

Just prior to the two paragraphs Sellars has cited in the quotation above (A736–7/B764–5), Kant had summed up his “Discipline of Pure Reason” section by stating that philosophy must curb its dogmatic pretensions and in “its transcendental efforts” bring reason “back to modest but thorough self-knowledge by means of a sufficient illumination [or clarification: Aufklärung] of our concepts” (A735/B736). The subsequent two paragraphs then make clear that philosophical cognition cannot consist in merely “analytic judgments” about given concepts, which “do not really teach us anything more about the object than what the concept that we have of it already contains in itself”; nor can philosophical cognition consist in “direct synthetic judgments from concepts.” Rather, the objective validity of synthetic judgments a priori is shown “only indirectly through the relation of these concepts to something entirely contingent, namely possible experience,” so that “an object of possible experience” is “presupposed.” So while a synthetic a priori proposition such as that of causality “must be proved, it is called a principle and not a theorem because it has the special property that it first makes possible its ground of proof, namely experience, and must always be presupposed in this” (A736–7/B764–5).

The above passages are consonant with Sellars’ analytic reading of Kant’s transcendental method as (1) presupposing that we do have empirical knowledge or cognition, which is expressible most generally in the concept of an object of possible experience, and (2) as consisting in an “illuminating” analysis of that complex concept, which reveals the necessary role of synthetic a priori principles as making such an experience possible. In this way Kant’s philosophical cognition could plausibly be described as consisting of “analytic knowledge about synthetic knowledge,” as we saw Sellars put it (KTE §10n2), since the analysis draws out the necessity of the relevant a priori syntheses for the possibility of experience. Furthermore, in the second of the two contexts in the Discipline cited above by Sellars, Kant indicates that every transcendental proposition “proceeds solely from one concept, and states the synthetic condition of the possibility of the object in accordance with this concept” (A787/B815). Kant explains that in the Transcendental Analytic, the (SAP) causal principle, for example, was drawn out (ausgezogen) “from the unique condition of the objective possibility of a concept of that which happens in general,” by showing that “the determination of an occurrence in time . . . as belonging to experience would otherwise be impossible” (A788/B816, italics added).

I want finally to suggest that Sellars’ various comments on the role of the transcendental unity of apperception (TUA) in the transcendental deduction bring out another crucial respect in which Kant’s transcendental method, on this reading, is plausibly taken to provide “analytic knowledge about synthetic knowledge,” again without assuming SAP principles in mathematics, natural science, or metaphysics as premises. And as I will also note further below, I believe that Sellars himself embraced Kant’s conception of the represented or thought-unity of the thinking self, developed in terms of his own broadly functionalist conceptual-role view of the nature of thought and intentionality, as a pivotal factor in his overall attempted “synoptic vision” or “fusion” of the “manifest and scientific images” of the human being in the world (cf. O’Shea 2016, Ch. 7)

Rosenberg 1986). I will focus here on the methodological aspects of Sellars’ account, which come to the fore straightaway when he remarks that while the TUA is perhaps “the central concept of the Critique, only recently has British and American philosophy freed itself sufficiently, first from its positivistic heritage, then from its anti-systematic bias, to be in a position to translate it into congenial terms” (I’ §6).

After allowing himself “to use the referring expression ‘I’ as a common noun,” the conceptual distinction involved in the TUA with which Sellars begins is “to the effect that ‘an I thinks of a manifold’ is not to be confused with ‘an I has a manifold of thoughts’ ” (1972 ‘I’ §§3–7, italics added). Whatever kind of thing an I might turn out to be, there is a distinction between a manifold (for example, a succession) of thoughts or experiences in a consciousness, and the single thought or experience of such a manifold as a manifold (for example, as a succession). Sellars’ Kant thus begins with our concept of a thinking self, not with any claim as to what the nature of such an I might be, as the most abstract form of our cognition or knowledge of a world about which we are not being ab initio sceptical, and the analysis of which ultimately reveals any such radical scepticism to be misconceived.

Kant’s categories or pure forms of thought are accordingly interpreted by Sellars as the “ways in which many thinkings constitute one thinking” (of course, why these particular ways is just one of the questions that arise in that context), and in particular as these concepts function in the “larger context” of our judgments about a world of objects and events in time and space (I’ §§8–10). Sellars then presents Kant’s TUA functioning in the latter, larger context as relating two TUA claims “analytically,” such that the a priori “form of empirical knowledge” takes the following form (the intervening comment below, “From this...”, is also Sellars):

An I thinking a complex spatial-temporal-causal system of states of affairs – including, say, $\alpha$ and $\beta$. (The synthetic unity of apperception).

From this synthetic unity, as Kant points out, it follows analytically that

The I which thinks $\alpha$ is identical with the I which thinks $\beta$. (The analytic unity of apperception).

Sellars presumably has in mind passages in the first two sections of the B-Deduction (§§15–16) such as this one:

Therefore it is only because I can combine a manifold of given representations in one consciousness that it is possible for me to represent the identity of the consciousness in these representations itself, i.e., the analytical unity of apperception is only possible under the presupposition of some synthetic one. The thought that these representations given in intuition all together belong to me means, accordingly, the same as that I [can] unite them in a self-consciousness. . . . (B133–4, boldface in original)

On the present understanding of Kant’s method, and in particular of its non-sceptical starting point of analyzing the concept of an object of our possible experience in general, the complex analytic entailment or conceptual presupposition in question can be drawn out in either of two
directions, as is in fact illustrated by the two passages above from Sellars and Kant. In this I agree with Sellars’ student Jay Rosenberg with respect to what he characterized as Kant’s conceptual “thesis of the mutuality of self and world” (1986, 6).

There is obviously a lot more to be said about these last claims as the basis for a reading of Kant’s transcendental deduction. However, I hope that part I has at least clarified and rendered plausible Sellars’ “analytic” interpretation of Kant’s transcendental method. I turn now in what remains to an examination of how Kant’s method, so interpreted, was further adapted and reformulated by Sellars, in light of the linguistic turn, in order to produce some of the currently most influential aspects of Sellars’ own philosophy. For present purposes I will present these Sellarsian themes compactly and in overview rather than in detail, in order to highlight and tie together their specifically Kantian methodological character.

II. The Kantian Methodological Roots of Sellarsian Themes on Mind, Knowledge, and Nature

From his first serious readings in philosophy in his early twenties as a Rhodes Scholar BA at Oxford (1932–34) onward, Sellars was convinced that Kant was essentially right about the necessarily holistic and epistemically anti-sceptical conceptual structure that constitutes our direct perceptual knowledge of one unitary spatiotemporal, objectively law-governed world of physical objects and events in nature, the methodological underpinnings of which have been sketched above. Sellars, however, unlike Kant, thought that this Kantian conceptual structure itself would somehow have to be philosophically naturalized, in the sense of providing an explanation of its own possibility as arising and functioning wholly within the scientifically conceived natural world. That is, for Sellars the “synoptic” philosopher must account not only for the origin but also for the irreducibly normative and social functioning of our Kantian conceptual-linguistic and sensory representational capacities consistently with an exhaustively scientific naturalist, evolutionary understanding of reality from top to bottom. The key to what I call Sellars’ reconciling Kantian naturalist project was to provide, as he put it in his 1974 “Autobiographical Remarks,” “an adequate naturalistic philosophy of mind,” where in particular what “was needed was a functional role theory of concepts which would make their role in reasoning, rather than a supposed origin in experience, their

---

14 An indication of this is that one can conduct the analysis in the transcendental deduction, as Kant himself does, both “starting from below” with the concept of an object of our successive apprehensions, and “starting above” with the transcendental unity of the “I think” in any experience (cf. O’Shea 2012, 132–49 for an introductory clarification of each direction of analysis in the Deduction).

15 Rosenberg 1986, 5–6: “What Kant argued, in his ‘Transcendental Deduction’, was that such an objective ‘synthetic unity of apperception’ was not a mere phenomenological given but was, in fact, correlative to the (subjective) ‘transcendental unity of apperception’. In other words, the conditions according to which an experienced world was constituted as an intelligible synthetic unity were at the same time the conditions by which an experiencing consciousness was itself constituted as a unitary self. That an experiencer represents the encountered world as categorially structured in space and time, Kant claimed to show, was a condition of the very possibility of his representing himself as a unitary subject of his world at all.” The “mutuality thesis” is an analysis of the conceptual structure of our cognitive intentionality, one which reveals, inter alia, the necessity of an objective conceptual a priori. It is not an attempt to prove deductively, from non-epistemic premises concerning subjective consciousness, that we have knowledge or objectivity.

primary feature. The influence of Kant,” he reports, “was to play a decisive role” in his own developing views on these matters (Sellars 1974, 285–6, italics added).

Also vital to this project of thoroughly integrating the Kantian-normative “logical space of reasons” within the scientifically natural world was Sellars’ famous rejection, in “Empiricism and the Philosophy of Mind” (1956, EPM), of what he called the myth of the given. In its most familiar form the latter myth, which Sellars argues is based on a variety of perennially tempting philosophical confusions, is the idea that there are some types of “immediate” sensory or rational cognition or states that by themselves — that is, without any essential dependence on the sorts of holistic conceptual structures discussed above — allegedly provide the experiencer or thinker with a direct recognition of what sort of quality or particular is being experienced (for example, this sensed redness), or of what basic facts or what self-evident principles of reason are directly apprehended or grasped (for example, that every event has some cause). Again, for Sellars, it was Kant’s account of the thoroughly conceptualized nature of our sense-perceptual cognition of objects, when integrated with insights from the pragmatists, Carnap, Wittgenstein, et al., on the importance of language in partly constituting the more mature forms of human conceptual cognition and reasoning, that provided the key to Sellars’ own understanding of these crucial matters.

We can see how these Kantian themes developed, for Sellars, out of the methodological considerations discussed in part I above. In “Toward a Theory of the Categories” (1970, TTC) Sellars explored what is involved in Kant’s view that “the pure categories are essential moments in the definition of the concept of an object of experience in general” (TTC §3). His analysis eventually brings out, once again, the constitutive necessity in experience of certain non-analytic principles as objectively valid prior to (that is, on conceptual grounds that are valid independently of) any particular courses of sense experience given a posteriori. For Sellars, Kant’s theory of categories represented a sophisticated advance in the tradition of Ockham’s earlier strategy of construing categories as meta-conceptual classifications of “mental words,” i.e., of conceptual items analogous to different types of linguistic category. In Kant this became “the idea that categories are the most generic functional classifications of the elements of judgments” (TTC §23). For example, the judgment that the sunlight warmed the stone employs an empirical causal concept (x warms y), and this has as its underlying logical form the ground/consequence hypothetical form of judgment that structures the category of causal dependence (if the sun shines, then the stone warms), here applied within experience. In Sellars’ own philosophy such categorial functional classifications of conceptual roles fall within a more comprehensive nominalist theory, building on insights from Carnap’s metalinguistic conception of ontology after the linguistic turn, of the meta-conceptual classificatory role not only of categorial concepts but of all abstract terms. Crudely put, philosophical vocabulary concerning, for example, properties, individuals (substances, particulars), and kinds are analyzed as culturally evolved ways of functionally classifying the socially rule-governed roles of predicates, names, and common nouns across different languages (including, for Sellars, ‘mentalese’ as a partly acquired ‘language of thought’) in our judgments about objects and events.

Kant’s focus on the logical forms of judgment in general, as the clue or guiding thread to his analysis of the most basic categorial forms involved in the concept of an object of cognition in general, was thus on Sellars’ view a crucial methodological advance. If we now consider the role of the categorial principles of understanding in their application to any finite sensory-perspectival experience, Sellars takes Kant’s analysis to have shown, correctly, that “the knowability of objects [must] consist, in large part, of inferential knowability,” and in particular “the concept of inferability
in accordance with laws of nature” (TTC §51).17 For example, in our spatiotemporal form of
experience, experience of an object or appearance here-now is an experience, as Sellars puts it, of
what is “on its way to being a there-then” object or appearance in the past, and “on its way from
having been a there-then” object or appearance in the future (TTC §53). So again, what “Kant takes
himself to have proved is that the concept of empirical knowledge involves the concept of
inferability in accordance with laws of nature. To grant that there is knowledge of the here and now
is . . . to grant that there are general truths of the sort captured by lawlike statements” (TTC §54;
cf. KTE §36).18 And Sellars once again concludes by tying these points back to the Kantian
methodological standpoint elaborated in part I above:

Thus, the transcendental knowledge that spatio-temporal objects of knowledge must
conform to certain generalizations which are themselves logically synthetic is itself, according to
Kant, analytic. (TTC §54; italics in original)

As Sellars goes on to clarify further in this context (TTC §55), such synthetic principles are thus
shown to be necessary relative to the assumption that we do have empirical knowledge of
spatiotemporal objects in general.

It is important here again to bear in mind the ways in which Sellars adapted these Kantian
methodological claims in “transcendental logic” about the necessity of inferential empirical
knowledge of laws of nature, to the post-linguistic turn setting of what he characterizes as
transcendental linguistics (KTE §40) or pure pragmatics (TTC §51). Put briskly, one largescale Sellarsian
adaptation of Kant’s argument for the necessity of logically synthetic yet necessary generalizations
or laws within any framework of empirical knowledge was introduced in order to take account of
the ever more striking instances since Kant’s time of radical conceptual change in logic, mathematics,
and science. Here Sellars was able to build upon C. I. Lewis’s attempt to incorporate Kant’s key
insight into his pragmatic (or framework-relative) conception of the a priori (Lewis 1929). Going beyond
Lewis, the epistemic requirement of lawful inferability was deepened empirically and semantically
by Sellars in his account of the necessity of normatively sanctioned patterns of material inference (as
opposed to purely formal-logical and analytic inference) to account for the very meaning or
conceptual content of any empirical term, however basic. Both of these broadly pragmatist
adaptations of Kant’s methodological insights and conclusions featured early on in Sellars’ work,

17 Obviously the relevant sense of inferability here is inferability within the objective domain, and not the
misconceived Cartesian or “problematic idealist” attempt to infer from the subjective to the objective (cf.
Kant’s Refutation of Idealism). Note also that Sellars is not rejecting non-inferential or direct perceptual
knowledge of physical objects: one can “recognize the essential and irreducible role of inference without
denying the existence of non-inferential knowledge of the here-now,” which in fact Sellars accepts in his own
Kantian account of our perceptual knowledge (TTC §51). Sellars repeatedly stresses that non-inferential
knowledge with inferential semantic and epistemic presuppositions is not the same as the Myth of the
(presuppositionless) Given that he famously rejects (1956 EPM I §1 and VIII §32). He also does not reject
the idea that something is “given” to our passive sensibility in sense experience, but of what sort depends on
Kant’s enduring insight that the categories are already involved even in our most basic singular sensible
intuitions themselves: “To be able to have intuitive representings, then, is to have all the conceptual
apparatus involved in representing oneself as acquiring empirical knowledge of a world one never made”
(KTE §36).

18 I have offered an elaboration of the grounds for this form of argument, as exhibited in a line of thinking
that runs from Kant through C. I. Lewis to Sellars and Brandom, in O’Shea 2015. So I will not pursue it
further here, though it is central to Sellars’ (and, for example, Brandom’s) Kantian methodological heritage.
most explicitly in his 1953 paper on Lewis entitled “Is There a Synthetic \textit{A Priori}?“ Both adaptations were also essential to Sellars’ famous arguments concerning the myth of the given, since no recognition or awareness of any item as being of a certain sort is given independently of those inferentially structured and (for Sellars) always revisable and replaceable conceptual contents that make it possible to recognize anything as being any sort of item in the first place. As Sellars expressed this key thesis in \textit{EPM}, in a way that clearly reflects his Kantian transcendental outlook:

\begin{quote}
For we now recognize that instead of coming to have a concept of something because we have noticed that sort of thing, to have the ability to notice a sort of thing is already to have the concept of that sort of thing, and cannot account for it. (\textit{EPM} §45; italics in original)
\end{quote}

Again, however, Sellars’ holistic \textit{concept-first} and anti-\textit{“concept-empiricist”} Kantian epistemology must be understood in the context of his fallibilist and explanatory pragmatism about conceptual change, and more broadly in light of his views on the socially norm-governed and scientifically evolving nature of our conceptual-linguistic capacities in general.

I noted earlier the driving \textit{scientific naturalist} dimension of Sellars’ fundamental philosophical aspirations. In fact, as Sellars saw things, it was ironically \textit{Kant’s} conceptual methodological insights – which had inspired and fortified Sellars’ own views about epistemological and semantic holism, the myth of the given, the normative-functionalist nature of thoughts and intentions (and hence, of persons), and the meta-classificatory, rule-ish nature of meanings and other abstract entities – that first \textit{made possible} a genuinely coherent and thoroughgoing \textit{scientific naturalism}, one that might truly account for, rather than explain away, the irreducibly normative nature of all of the various human phenomena just mentioned. For as was noted briefly in the closing paragraphs of part I, Kant’s conception of a \textit{thinker} (an I, or a self) as a “form of representation” (A345–6/B404) rather than the representation of a certain kind of \textit{thing},\textsuperscript{19} whether material or immaterial, was for Sellars (though he recognizes, not for Kant himself) the key to a normative-functionalist yet exhaustively naturalistic synoptic vision of the ultimate nature of thoughts as fully materially realized in the environmentally engaged states of embodied persons.\textsuperscript{20}

Furthermore, the important ontological effect of the exposure of the myth of the given, for Sellars, was thereby to expose any conceptual framework, no matter how basic, as in principle \textit{replaceable} by any framework that would more adequately conceptualize and explain the relevant empirical phenomena. On Sellars’ own synoptic vision, whatever its ultimate merits (and it is not

\textsuperscript{19} As Kant puts it in the \textit{Paralogisms}: “At the ground of this doctrine we can place nothing but the simple and in content for itself wholly empty representation \textbf{I}, of which one cannot even say that it is a concept, but a mere consciousness that accompanies every concept”; for this “consciousness in itself is not even a representation distinguishing a particular object, but rather \textit{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” (A345–6/B404, italics added) Roughly put, the thinking self is a unity in virtue of the \textit{a priori} forms of conceptualized unity in terms of which it necessarily represents an empirical \textit{world of objects} that exists independently of its experiences. Again, for further interpretive details on this point from the present perspective, see Sellars 1972, Rosenberg 1986, 2005, and O’Shea 2012, Ch. 4.

\textsuperscript{20} Note, also, that I have left entirely aside for present purposes Sellars’ grappling with the problem of qualitative sensory consciousness, or the “hard problem” as it is called today. There is of course a Kantian dimension to this aspect of Sellars’ philosophy, by way of Kant’s views on \textit{sensibility} (cf. Sellars 1968, Ch. I).
easy to make Sellars’ own systematic vision hang together in its entirety), this ongoing process of conceptual change must ultimately include the explanatory reconception of the basic object-ontology of the sophisticated common sense or “manifest image” of the perceived world itself, resulting ideally in its wholesale replacement by a more adequate scientific process-ontology. My concluding suggestion, however, is that the most important and overriding lesson that Sellars himself drew from Kant’s transcendental methodological analyses was that even such an imagined reduction and replacement as he envisions of our common sense ontology of objects by an ideal scientific image of the world would not be an elimination or reduction of persons as conceptually thinking, intentional agents perennially engaged in reconceiving and practically transforming both their shared social worlds and the independent natural world they inhabit and conceptually represent.22

**BIBLIOGRAPHY**


---

21 For recent assessments of Sellars’ famously divided philosophical legacy in general, which has unfortunately but not unreasonably ossified into an often misleading distinction between “left-wing vs. right-wing Sellarsians,” see all the contributions to the Sellars and his Legacy volume (O’Shea, ed. 2016).

22 I am very grateful to two BJHP referees for their helpful queries and insightful requests for further clarification, and my thanks also to Gabriele Gava and to Michael Beaney.


Sellars, Wilfrid. 1972 (‘I’). ‘…this I or he or it (the thing) which thinks.’ The presidential address, American Philosophical Association (Eastern Division) for 1970, *Proceedings of the American Philosophical Association* 44: 5–31; in 2002a (KTM), 341–62.


