S T R A W S O N ’ S  M E T A C R I T I Q U E

Anil Gomes
Trinity College, University of Oxford

Forthcoming in A. Bengtson, B. De Mesel and S. Heyndels (eds.) P.F. Strawson and His Legacy, Oxford University Press [submitted 2022]

What is the status of the claims which make up Kant’s arguments in the Critique of Pure Reason? This question seemed to Kant’s contemporaries to require a metacritique. Strawson’s criticisms of Kant should be understood in this context: as raising a metacritical challenge about Kant’s grounds for the claims which make up his arguments. What about the claims which make up Strawson’s own arguments in The Bounds of Sense? I argue in this chapter, against what I take to be the general consensus, that Strawson did not and should not have understood these claims to be analytic. Rather he is somewhat puzzlingly committed to our possessing non-analytic but still a priori knowledge of his claims. What could such knowledge consist in? I’ll extract from G.E. Moore’s early writings on Kant a model for understanding such knowledge, one which enables us to better appreciate the way in which Strawson’s methodology dovetails with Kant’s own.

1. Introduction

Strawson’s Kant occupies an equivocal place in the philosophical study of Kant. On the one hand it is charged with revitalising interest in Kant, opening ‘the way to a reception of Kant’s philosophy by analytic philosophers’ in the
words of Hilary Putnam (1998: 273). On the other, it is criticised by Kant scholars for its less than comprehensive engagement with Kant’s works and its less than charitable interpretation of his theses. ‘I have not been assiduous in studying the writings of Kant’s lesser predecessors, his own minor works or the very numerous commentaries which two succeeding centuries have produced’, Strawson warns us at the start of The Bounds of Sense (1966: 11). This, says Lucy Allais, ‘seems an enormous understatement’ (2016: 893). And yet there is no denying the power and creative verve of Strawson’s philosophical interrogation of Kant’s text.

Strawson’s interest in Kant arose from the peculiarities of Oxford’s degree in Philosophy, Politics and Economics. At the time he was a student, there were two special subjects which those who wished to specialise in philosophy were obliged to take: Logic and Kant. The latter involved the study of two texts: the Critique of Pure Reason and the Groundwork of the Metaphysics of Morals. Strawson found the Groundwork ‘deeply impressive’ but thought it conceived its subject altogether too narrowly. In the Critique of Pure Reason, on the other hand, he found ‘a depth, a range, a boldness, and a power unlike anything I had previously encountered’ (Strawson, 2003: 8). Kant’s influence threads through Strawson’s writings, not least in both the conception and conclusions of Individuals (1959), a work ‘subtly and in part consciously influenced by [the first Critique]’ (Strawson, 2003: 8). But it is in the book which grew out of his lectures on the first Critique—The Bounds of Sense—that Strawson takes on Kant explicitly.1

Strawson’s central approach to the first Critique is captured by the three-fold pun which constitutes his title. It is an echo of a title that Kant considered for the first Critique.2 And it plays on the ambiguity of the word ‘sense’ which can denote both sense-experience and sense-meaning. It thus, Strawson tells

---

1 Strawson continued to publish on Kant’s theoretical philosophy, most notably the four essays collected together in his Entity and Identity (1997). These are important for tracing the development of his views.

2 In a letter to Marcus Herz from 1771, Kant writes ‘I am therefore now busy on a work which I call “The Bounds of Sensibility and of Reason”’ (Correspondence 10:123); he repeats the title in a letter to Herz from 1772, writing of a work which might perhaps have the title, The [Bounds] of Sensibility and Reason’ (Correspondence 10:129). Zweig’s translation renders the second title as The Limits of Sensibility and Reason, but both titles use the same German term, ‘Grenzen’.
us in his Preface, ‘alludes compendiously to the three main strands in [Kant’s]
thought’ (1966: 11). First, that there is a lower limit on sense, ‘a certain
minimum structure is essential to any conception of experience which we can
make truly intelligible to ourselves’ (1966: 11). Second, that there is an upper
limit on sense, for ‘the attempt to extend beyond the limits of experience the
use of structural concepts, or of any other concepts, leads only to claims
empty of meaning’ (1966: 11–12). Finally, that Kant draws these limits from
‘a point outside [the bounds of sense], a point which, if they are rightly drawn,
cannot exist’ (1966: 12). Strawson’s project is to extract what is valuable in the
first two strands from what is supposed to be the incoherence of the last.

Strawson’s negative assessment of this final strand drew immediate discussion.
It includes his rejection of Kant’s metaphysics of transcendental idealism and
his rejection of the ‘imaginary subject of transcendental psychology’ (1966:
32). The response to the rejection of transcendental idealism was not to
defend the doctrine so understood against Strawson’s attack. It was rather to
deny the attribution of that doctrine to Kant. Graham Bird had already
offered an alternative in his Kant’s Theory of Knowledge (1962) and Henry
Allison’s Kant’s Transcendental Idealism (1983) explicitly used Strawson as a
foil for his own supposedly more sympathetic interpretation. The response of
these defenders was that Strawson’s account of transcendental idealism was
exegetically unsound but philosophically on point.

The response to Strawson’s rejection of transcendental psychology made a
contrasting case. It did not deny that Kant engaged in such a subject. It rather
denied that it was problematic for him to do so. This is most clear in Patricia
Kitcher’s Kant’s Transcendental Psychology (1990) which defended
transcendental psychology as neither imaginary nor excisable from Kant’s text.
Here the thought was that Strawson is right to understand Kant as engaged
in transcendental psychology but wrong to think there is anything improper
in being so.

Strawson’s discussion of transcendental idealism has provoked reams of
commentary. My focus in this essay is his criticism of the ‘imaginary subject
of transcendental psychology’ (1966: 32). It is a quotable and oft-quoted line,
characteristic of Strawson’s ear for the right phrase. But what exactly is the
nature of Strawson’s criticism? That is the topic of this essay. It is a small
exegetical issue which turns out to have important implications for understanding what Strawson thought he was doing when he was doing philosophy. And perhaps also for what Kant thought he was doing as well.

Here is the structure of what follows. In §2 I’ll reconstruct Strawson’s criticism of transcendental psychology and suggest that it instantiates a challenge which stretches back to the very first readers of the Critique of Pure Reason—a challenge about Kant’s grounds for the claims which make up the argument of the Transcendental Deduction. Does Strawson’s own reconstruction of the Deduction avoid this problem? Answering this question requires us to understand how Strawson understood the grounds for the claims which make up his own version of the Deduction. In §3 I shall argue, against what I take to be the general consensus, that Strawson did not and should not have understood the claims which make up his own argument to be analytic. Rather he is puzzlingly committed to our possessing non-analytic but still a priori knowledge of the claims which constitute his own argument. But what could such knowledge consist in?

In the second part of the essay, I’ll sketch an answer to this question, starting in §4 with a discussion of G.E. Moore’s early writings on Kant and a consideration of Moore’s own methodology. Moore too was committed, I’ll suggest, to our possessing non-analytic but still a priori knowledge of the claims which constitute his own arguments. Consideration of Moore’s methodology will offer a model in §5 for understanding what Strawson is up to in his version of the Transcendental Deduction. And it will allow us in §6 to better appreciate the way in which Strawson’s methodology dovetails with Kant’s own.

2. Kant (I)

Why does Strawson think that transcendental psychology is an imaginary subject? His use of that phrase occurs in his overview of Kant’s Transcendental Deduction of the Categories. It is worth quoting the passage in full:

[The Transcendental Deduction] is also an essay in the imaginary subject of transcendental psychology. Since Kant regards the necessary unity and connectedness of experience as being, like all
transcendental necessities, the product of the mind’s operations, he feels himself obliged to give some account of those operations. Such an account is obtained by thinking of the necessary unity of experience as produced by our faculties (specifically by memory and imagination controlled by understanding) out of impressions or data of sense themselves unconnected and separate; and this process of producing unity is called by Kant “synthesis”. The theory of synthesis, like any essay in transcendental psychology, is exposed to the ad hominem objection that we can claim no empirical knowledge of its truth; for this would be to claim empirical knowledge of the occurrence of that which is held to be the antecedent condition of empirical knowledge. (32)

The Transcendental Deduction aims to show that we are entitled to use a set of pure concepts: the categories. It does this by arguing that these concepts are a priori conditions on the possibility of experience and that this suffices to show how they can relate to objects a priori and why we are entitled to use them. Strawson’s gloss on this argument reads Kant as showing that the categories are conditions on the possibility of experience by identifying features of experience—its necessary unity and connectedness—which are imposed by our faculties through a process of synthesis. The imposition of this unity through synthesis is supposed to explain why the categories are conditions on the possibility of experience.

Strawson claims that this aspect of Kant’s argument—the claim that the mind imposes unity through the operation of synthesis—is exposed to ‘the ad hominem objection that we can claim no empirical knowledge of its truth’ (32). It is not clear what makes this objection ad hominem but it at least seems true. Indeed, its truth looks overdetermined so far as Kant is concerned. Kant holds that empirical knowledge is established by experience and that experience can only tell us what is, never that it must be so (B3). It follows that experience cannot be a source of knowledge, of necessary truths, that they are necessary. So if it is known to be a necessary truth that the mind imposes unity in experience through the operation of a process of synthesis, then experience is not in a position to support the claim and it cannot be an object of empirical knowledge.⁵

⁵ On the question of whether it is a necessary truth, see (Gomes, Moore, Stephenson 2022) for relevant discussion.
In fact, we do not even need to assume that the claim is a necessary truth. For Kant also holds that experience can only tell us about how things appear. So if the claim that the mind imposes unity through synthesis were an item of empirical knowledge, it would be a claim only about how things appear. But a truth about how things appear cannot itself be a condition on the *possibility* of appearances. So the mere fact that this claim is supposed to be a necessary condition on experience shows that we cannot have empirical knowledge of its truth.

Note that Strawson’s charge here is independent of his characterisation of the cognitive faculty as *producing* some aspects of our experience. This terminology reflects Strawson’s reading of Kant as endorsing ‘a relatively familiar kind of phenomenalistic idealism’ (240) on which the objects of experience are constructed out of sense-impressions. Some have thus thought that the objection he raises to transcendental psychology is inseparable from his interpretation of transcendental idealism more widely such that the rejection of transcendental psychology stands or falls with the rejection of transcendental idealism. But this is not the case. Say one thought only that synthesis is one of the ‘epistemic conditions’ under which objects can be known by us (Allison 1983: 10). Strawson can still hold that this claim is exposed to the objection that we can claim no empirical knowledge of its truth.

If this is right, then the claims which make up Kant’s transcendental psychology cannot be empirical truths. How do we get from there to the claim that transcendental psychology is an imaginary subject? Strawson offers us no more in this initial discussion but further considerations are forthcoming in the discussion of synthesis in §6 of the discussion of Objectivity and Unity in Part II of the book. This is the part in which Strawson lays out his own version of the Transcendental Deduction of the Categories. Strawson there repeats the charge that the doctrine of synthesis cannot be a matter of empirical knowledge, this time emphasising the reason that it is not and cannot be an object of empirical self-consciousness.

4 E.g. (Guyer 2017: 368)
Strawson then goes further and considers whether there is some other way in which we might know that the mind imposes unity through synthesis. He considers those passages in the Deduction in which Kant seems to suggest that we are conscious of the activity of synthesis. And Strawson thinks they are to be interpreted away:

Kant does not, after all, think that we have a special kind of experience or awareness of the self or its activity, distinct from that empirical self-consciousness in which, as he holds, we are aware only of appearances of ourselves. (95)

If right, this rules out an alternative way of our coming to know that the mind imposes its unity through synthesis—through some special non-empirical awareness of the activity of synthesis. But it does not yet show that there is no way of coming to know the claim. And that is what we need if Strawson is to show that the subject is imaginary.

I think there are two assumptions lying behind Strawson’s discussion here and once we make them explicit we can see both why the challenge which he presents to Kant is a good one and how it instantiates a schema of objection which has been raised since the earliest engagements with the first Critique. Note, first, a curious omission. Strawson does not here consider whether the doctrine of synthesis might be an analytic truth, knowable in whatever way we know analytic truths more generally. Perhaps Strawson thinks it obvious that it is not an analytic truth or that Kant could not have thought it such. This is the first assumption. One might support it by noting that it does not seem built into the concept of experience that its unity is the result of a process of synthesis. Or alternatively, that Kant’s aim in the Transcendental Deduction of the Categories is to show the objective validity of the categories and that this requires showing that their objects are really possible. But analytic truths determine only logical possibilities. So if the doctrine of synthesis were an analytic truth, it would need supplementation with non-

5 Strawson cites B133 and A108; he might further have appealed to the footnote at B157 and its suggestive comments about the relation of the ‘I think’ to activity.
analytic considerations to show the real possibility of the categories. Strawson must be assuming that this cannot be done.⁶

Given this assumption, Strawson may conclude that the doctrine of synthesis is not an item of empirical knowledge nor an item of analytic knowledge. In order to complete the argument, he needs further to show that we cannot have non-analytic a priori knowledge of the doctrine of synthesis. This, I suggest, is how we should understand his rejection of ‘a special kind of experience or awareness of the self or its activity, distinct from… empirical self-consciousness’ (95). Knowledge based on such special experience or awareness would be non-analytic, in virtue of its basis in experience or awareness. But it would also be a priori, in virtue of being distinct from empirical self-consciousness. This is Strawson’s second assumption: that such special experience is the only way in which Kant could claim non-analytic a priori knowledge of the doctrine of synthesis.

We can now see the shape of Strawson’s challenge to the subject of transcendental psychology. First, he claims that we cannot have empirical knowledge of the doctrine of synthesis. Second, he assumes that we cannot have analytic knowledge of the doctrine of synthesis. Finally, he claims that we cannot have a special kind of awareness of synthesis and assumes that this would be the only way to have non-analytic a priori knowledge of the doctrine of synthesis. Since these exhaust our ways of knowing, the doctrine of synthesis cannot be known: it is part of the imaginary subject of transcendental psychology.

Once set out in this way, Strawson’s challenge instantiates a more general form of objection to the first Critique, one which traces back to its very first interlocutors. The Critique of Pure Reason aims to explain the possibility of synthetic a priori judgment. Kant’s explanation of this possibility involves certain claims about the structure of the mind. There are two faculties to the cognitive mind, a passive faculty of sensibility and an active faculty of the understanding. Neither faculty can be reduced to the other. But they are individually necessary and jointly sufficient in finite beings for knowledge

⁶ See (Gomes, Moore, Stephenson 2022: §6.1) for discussion relevant to the question of whether the doctrine of synthesis is an analytic truth.
Each of these faculties has its own representations by means of which it relates to objects. Objects are given to us in sensibility by means of intuitions but we think of objects through the understanding by means of concepts. And Kant’s explanation of the possibility of synthetic a priori knowledge turns on the claim that each of these faculties has an a priori element. Sensibility has pure intuitions, space and time; the understanding has pure concepts, the categories. It is these a priori elements to sensibility and the understanding—their pure forms—which explains the possibility of synthetic a priori knowledge.

What is the status of these claims about the structure of our cognitive faculties and their role in producing experience? Strawson’s version of the challenge focuses on the doctrine of synthesis specifically but the question ranges more widely. It is a question about the claims about the structure of our cognitive faculties which make up the first *Critique* rather than the claims which are established on their basis. That some account is owed of how we know these claims was recognised by Kant’s contemporaries. His sometime friend and close interlocutor Johann Georg Hamann wrote in 1784 that they show the need for a *metacrítica*: an examination of the foundations and fundamentals of philosophical critique itself. Hamann sent his ideas to Herder who sent them on to Jacobi and through them they set the foundations for the way that the post-Kantian German idealists engaged with Kant’s text. In its most general form, the metacritical challenge to Kant is to explain how we know those claims which he appeals to in his explanation of the possibility of synthetic a priori knowledge.

7 Kemp Smith’s translation—used by Strawson—renders this term as ‘knowledge’. But there are textual and philosophical grounds for disquiet. It elides Kant’s distinction between *Erkenntnis* and *Wissen* if ‘knowledge’ is used to translate both terms. And it overemphasises the connections between *Erkenntnis* and the kind of propositional knowledge which has been the subject of much contemporary epistemology. Recent translations prefer the term ‘cognition’. I will continue to use the term ‘knowledge’ in order to remain connected to Strawson’s discussion but will flag any occasions where the translation is relevant. See (Gomes & Stephenson, 2016; Watkins & Willaschek, 2020; Schafer, forthcoming) for discussion of the issue.

8 In his *Metacrítica of the Purism of Reason*, reprinted in (Hamann 2007)

9 See (Beiser 1987: 37-43)

10 There is a version of the metacritical trilemma in Bennett (1966:16-17). Colin Marshall (2014) provides a comprehensive overview of the trilemma and the possible responses. Note
This gives us a framework for understanding Strawson’s objection to the imaginary subject of transcendental psychology. Once extended in the way I have suggested, it becomes a recognisable objection from the history of engagement with Kant’s first *Critique*. We will not consider here whether Kant succumbs to the challenge. Instead, I want to pursue a more *ad hominem* route. Does Strawson’s own version of the Transcendental Deduction avoid the metacritical problem which (I have suggested) he thinks afflicts Kant’s own? This is our concern in the rest of the essay.

3. **Strawson (I)**

Strawson’s challenge to Kant focuses on the doctrine of synthesis. And since Strawson makes no use of that doctrine in his reconstruction of the Deduction, the specifics of his objections get no purchase on the reconstructed argument. Our question then is whether the more general metacritical challenge applies to Strawson’s reconstruction. How do we know the claims which make up Strawson’s argument? This is the metacritical challenge to Strawson.

Answering that question requires a quick sketch of Strawson’s reasoning in the section Objectivity and Unity in Part II of *The Bounds of Sense*. I have argued elsewhere that the argument can be represented as follows:

1. The self-ascription of experiences requires possession of the concept of experience.
3. Possession of a conception of objectivity requires experience of objective things.\(^\text{11}\)

that nothing here turns on the use of Kemp Smith’s term ‘knowledge’ as a translation for ‘Erkenntnis’. The assumption is only that Kant’s claims about the structure of our cognitive faculty must have some positive epistemic status in order to explain the possibility of synthetic a priori cognition.

\(^{11}\) See (Gomes 2016) for elaboration and defence.
This beguiling argument was influential on a generation of philosophers and has been the impetus to many fruitful discussions.\textsuperscript{12} We can prescind from the details for our purposes. How does Strawson think we know the claims which make up his argument?

Many will think the answer obvious: through conceptual analysis. Strawson after all characterises his argument as one ‘which proceeds by analysis of the concept of experience in general’ (31), as one which establishes ‘a direct analytical connexion between the unity of consciousness and the unified objectivity of the world of our experience’ (96, my emphases in both quotations). Hans-Johann Glock stands for many when he characterises Strawson, in his editorial introduction to the collection Strawson and Kant (2003), as ‘the leading proponent of analytic Kantianism’ (Glock 2003a: 1), that view being one on which ‘the central insight of the Critique is an analysis of complex connections between concepts such as experience, self-consciousness, objectivity, space, time and causation’ (Glock 2003b: 30).\textsuperscript{13} This much is unproblematic. The interesting question is whether it follows that Strawson thinks that the claims which make up his argument are instances of analytic knowledge. That is invariably assumed.\textsuperscript{14} But there are reasons to be wary.

The first concerns Strawson’s discussion of Jonathan Bennett’s book, Kant’s Analytic (1966). It is Bennett who most clearly identifies both the claims which make up Kant’s arguments and their conclusions as analytic. ‘[T]he most interesting truths which Kant calls synthetic and a priori’, he writes, ‘are unobvious analytic truths’ (1966: 42). These are, to a rough approximation, those which are established by a series of steps involving only obviously analytic truths, such that the resulting truth is both analytic, in virtue of being established solely on the basis of analytic truths, and yet unobviously so, in virtue of the length of reasoning involved in establishing it (1966: 7-8). Others who have taken this line include Ralph Walker (1978: 18-19) and Richard Rorty who, in an posthumously published paper dating from the

\textsuperscript{12} These include (Bennett 1968; Harrison 1970; Rorty 1970; Walker, 1978: 116-121; Hurley 1994; Van Cleve 1999: 98-104). Its influence can be seen on (Evans 1982; McDowell 1994; Cassam 1997; Campbell 2002).

\textsuperscript{13} For similar characterisation of Strawson, see (Rorty 1970; Hacker 2003; Guyer 2017).

\textsuperscript{14} E.g. by Marshall (2014: 557) and Guyer (2017: 368)
1960s, writes ‘The only possible conclusion is that all of Kant’s remarks about human knowledge, must, on Kant’s own grounds, be construed as analytic propositions’ (Rorty 2020: 53). Bennett thinks the same is true of Strawson’s claims in Individuals (1966: 41-2).

Kant’s Analytic and The Bounds of Sense were published in the same year. Both instantiate a distinctive approach to the engagement of Kant’s texts. These genuine commonalities have allowed people to read into Strawson something like Bennett’s notion of the unobviously analytic. But Strawson’s review of Kant’s Analytic (Strawson 1968) suggests that this is a mistake. (Is it mischievous to say that those who lump Strawson and Bennett together have not been assiduous in studying Strawson’s own minor works?)

In that review, Strawson points out a series of problems with Bennett’s account of the unobviously analytic. Most basically, the claims which Bennett takes to be unobviously analytic simply do not fit the model of being established by a series of steps involving obviously analytic truths (1968: 334). Nor do the claims which make up Strawson’s version of the Transcendental Deduction. Either, then, they are not unobviously analytic or we need some alternative way of characterising the notion. But Strawson also expresses scepticism about ‘the utility of invoking the notion [of analyticity] to preserve the respectability of our metaphysics (1968: 335). This suggests that he does not take the claims in his reconstruction to be analytic.

The second reason for hesitation in classifying Strawson’s claims as analytic concerns his elusive remarks on the notion of the synthetic a priori. Say that the claims which make up Strawson’s argument are analytic. Then the claims which they support will be analytic as well. And there will be no use for the notion of the synthetic a priori in characterising such claims. This is the line that Bennett and Walker take (Bennett 1966: 42; Walker 1978: 18). Bennett, for instances, replaces the class of the synthetic a priori with the class of the unobviously analytic. Strawson is more cagey. For although he thinks that Kant ‘has no clear and general conception of the synthetic a priori at all’ (43) he does think that the notion can be used to pick out a class of propositions that have a distinctive character or status: those which are descriptive of the ‘fundamental general structure of any conception of experience such as we can make intelligible to ourselves’ (44). Strawson does not say that these
propositions are a special class of analytic propositions. But this is what we should expect him to say if he thought the claims which constitute his argument were analytic.¹⁵

These considerations are not decisive. Many readers will struggle to hear a difference between those claims which are descriptive of the fundamental general structure of any conception of experience such as we can make intelligible to ourselves and those claims which are analytically obtained through articulation of the concept of experience. And given that Strawson talks freely and easily of analysis and analytical (but not analytic) connections, the position I am ascribing to Strawson seems to be one on which analysis need not result in analytic knowledge. I do think Strawson is committed to such a claim. And I’ll give more evidence for this in a moment, if not quite from the horse’s mouth, then at least from the mouth of a horse in the same stable. For now, let us examine which options are open to Strawson if he denies that the claims that make up his arguments are analytic truths.

There are two. Either Strawson rejects the terms in which the metacritical trilemma is posed and does not have to choose between options. Or the conceptual analysis which delivers knowledge of the main conditions in Strawson’s argument is not supposed to result in analytic knowledge but one of either empirical knowledge or non-analytic a priori knowledge. Let us consider them in turn.

First, the rejection of the terms. There are a number of places where it looks like Strawson might reject the terms in which the metacritical trilemma has been phrased. In the review of Bennett, for instance, he does not forswear the notion of analyticity but instead expresses doubt about whether it is of any use in characterising Kant’s arguments (1968: 335). More generally, he expresses scepticism about whether some more basic explanation can be given of those claims that are descriptive of the fundamental general structure of any conception of experience such as we can make intelligible to ourselves. Kant took his transcendental idealism to be the explanation for the possibility of synthetic a priori knowledge. Strawson ‘see[s] no reason why any high doctrine should be necessary here’ (44). And that can suggest that far from

¹⁵ See (Stroud 2003) for an excellent discussion of Strawson’s notion of the synthetic a priori.
offering an answer to the metacritical challenge, he denies the obligation of answering its question.

There are some problems with taking this route on Strawson’s behalf. First, it is not clear how the metacritical challenge can be sidestepped. The challenge asks only what grounds one has to endorse the claims which make up one’s philosophical arguments. Of course, arguments have to begin somewhere. (‘What about the premises?’, Cian Dorr once asked, ‘Where did they come from? The premise factory?’ (Dorr 2010)). But it would be an odd philosopher who thought their premises lacking in any positive epistemic status. Second, Stawson’s remarks on the synthetic a priori do not themselves show that he resists the terms in which the metacritical challenge is put. He says only that an explanation of those fundamentally descriptive claims will not appeal to high theory. That is compatible with it appealing to something else. Finally, and most relevant dialectically, a defence of Strawson which has him resisting the terms of the metacritical challenge blunts his use of a version of that trilemma as an objection to Kant.

Turn instead to the second option. If Strawson does not reject the terms in which the trilemma is posed, and if he does not take the claims which constitute his own argument to be analytic, then that leaves him with two options: either the claims which constitute his own argument are known empirically or they are known in some non-analytic yet a priori way. The first of these options is unpromising. Strawson’s use of the term ‘analysis’ to pick out the activity in which he engaged is partly aimed to dissociate it from the kind of empirical support gained through a study of the mechanisms of self-conscious thought and experience. And, like Kant, Strawson is sceptical about whether experience alone could identify the kinds of necessary conditions which his version of the Transcendental Deduction sets out to establish.

What about the second disjunct? Could Strawson answer the metacritical challenge by holding that the claims which constitute his arguments are instances of non-analytic a priori knowledge? This option was supposed to be

16 See his remarks on the relation between Kant’s doctrine of synthesis and scientific investigation in (Strawson 1989).
a problem for Kant because it seems to entail that Kant's explanation of synthetic a priori knowledge appeals, in part, to instances of synthetic a priori knowledge. And that circularity looks to undermine the cogency of Kant's explanation. Since Strawson does not share Kant's notion of the synthetic a priori and since his aim is not to explain the possibility of such knowledge, there is no obvious circularity in his appealing to non-analytic a priori claims in the identification of that general structure of any conception of experience such that we can make intelligible to ourselves. So, on the face of it, nothing precludes Strawson from taking this route.

In fact, we can go one better. Not only is Strawson not precluded from taking this option, we have testimony for his endorsing it. Quassim Cassam, on the 50th anniversary of the publication of The Bounds of Sense, writes:

Strawson was absolutely clear in discussion that he never regarded experience of mind-independent objects as an analytically necessary condition for reflective experience and that many of his claims about necessary conditions of experience in The Bounds of Sense had a different status. He sometimes described them, somewhat mysteriously, as non-analytically but still conceptually necessary conditions. (2016: 915)

Non-analytic but still conceptually necessary. This is our third option. It shows that Strawson's response to the metacritical challenge would be to hold that the claims which constitute his reconstruction of the Transcendental Deduction are not known empirically, nor known analytically. They are non-analytic but still a priori truths.

Let us take stock. We are considering how Strawson would fare against his own metacritical challenge. I have suggested, conventional wisdom notwithstanding, that Strawson did not think of the claims which make up his arguments as analytic truths. Not only is this suggested by his review of Bennett and his account of the synthetic a priori, it is, according to Cassam, how Strawson himself described his claims. So despite the fact that The Bounds of Sense is involved in an analysis of any conception of experience that we can make intelligible to ourselves, the resulting claims are to be thought of as non-analytic but still conceptual, a priori truths. Cassam calls this
mysterious. And so it is. My aim in the second part of this essay is to see if we can make it less so.  

4. Moore

I noted above that the metacritical trilemma is as old as the first Critique itself. One striking place where it surfaces is at the very foundation of analytic philosophy. It is common, perhaps too common, to present analytic philosophy as arising from the rejection of neo-Hegelian idealism undertaken by Moore and Russell in Cambridge at the start of the twentieth century.  

One of the key moments in that process is Moore's transition from idealism to realism, a transition charted in the 1897 and 1898 dissertations which Moore submitted for the Trinity College Prize Fellowship examinations. At the start of the 1897 dissertation, Moore is an idealist of the McTaggart variety. By the end of the 1898 dissertation, he is the full-fledged realist that we know from the famous 1903 papers. Both dissertations are on Kant. Moore's central objection to Kant, repeated in various ways across the dissertations, is that the necessary conditions which Kant sets out are either purely psychological conditions about the way we are conditioned to think and behave, or else purely logical truths which cannot explain the possibility of the states and activities in question. We can think of this charge, in very broad terms, as challenging the claims about the cognitive faculties which constitute premises in Kant's arguments in the first Critique. Either these claims are analytic a priori truths, Moore says, in which case they cannot explain the possibility of synthetic a priori knowledge. Or they are empirical truths, in which case they concern only the psychological structure of how human beings actually operate and cannot be used to explain synthetic a priori knowledge. In 'Kant's Idealism', presented to the Aristotelian Society in

17 Wittgenstein's appeal to grammar has sometimes been understood as allowing for a kind of non-analytic, conceptual truth. Hacker (1996: 177-8) offers this notion as a salve to Strawson but does not see that something similar might already be operative in Strawson's text.
18 See (Bell 1999; Kalderon & Travis 2013; Gomes 2017a) for discussion of the way that the traditional story obscures key parts of that history
19 Recently published as (Moore 2011).
20 See (Moore, 2011: 141ff.) for a version of the dilemma in the 1898 dissertation and Caird's examiner's report (2011: 103) for a discussion of the issue as it arises in the 1897 discussion. The pages which follow give Caird's defence of Kant. The editors' introductory material provides further discussion (2011: xlvi-lxvii).
1903, Moore completes the trilemma by noting that if they are synthetic and a priori then ‘Kant has not, in his own words, “explained the possibility of all synthetic propositions a priori”’ (Moore 1903: 133).

Moore’s challenge to Kant, then—a challenge which forms part of the foundational texts of analytic philosophy itself—is an instance of the metacritical trilemma. Does this trilemma have force against Moore himself? These early dissertations do not contain a detailed account of methodology but the notion of analysis is already central to how Moore understands his approach. ‘[A] thing becomes intelligible first’, he tells us ‘when it is analysed into its constituent concepts’ (1898 dissertation, 2011: 168, my emphasis). Yet this coexists with a dismissive stance on analytic truth (1898 dissertation, 2011: 139-141). As his editors put it, ‘Moore does not think that philosophical analysis brings with it a commitment to regarding a priori philosophical truth as “analytic”’ (2011: lvi). This suggests the kind of distinction between analysis and analytic truths which I have proposed is found in Strawson.

How could analysis not result in analytic truths? The notion of analysis remains central in Moore’s mature writings, coming to signify, as John Wisdom wrote in 1931, a method that ‘Wittgenstein has lately preached and Moore long practised’ (1931: 195n.2). But what does such a method involve? We can extract an answer from the influential ‘A Defence of Common Sense’ (1925). This was Moore’s contribution to a series of invited essays by British philosophers which aimed ‘to give the contributors an opportunity of stating authentically what they regard as the main problem of philosophy and what they have endeavoured to make central in their own speculation upon it’ (Muirhead 1924: 10, quoted in Baldwin 2013). Moore focuses his essay on the importance of identifying the correct analysis of common sense propositions. How does such analysis work?

We get an answer in the final part of this paper, where Moore analyses the proposition ‘This is a hand’. He argues that analysis shows the existence of a sense-datum which is the subject of the proposition and thus that perception

21 See (Baldwin 2013) on ‘The Cambridge School of Analysis’.
involves a relation to sense data. How does analysis do this? Moore’s answer is instructive:

[I]n order to point out to the reader what sort of things I mean by sense-data, I need only ask him to look at his own right hand. If he does this he will be able to pick out something (and, unless he is seeing double, only one thing)... Things of the sort (in a certain respect) of which this thing is, which he sees in looking at his hand, and with regard to which he can understand how some philosophers should have supposed it to be the part of the surface which he is seeing, while others have supposed that it can’t be, are what I mean by ‘sense-data.’ (1925: 54).

What should we say about Moore’s philosophical method here? It offers us, in effect, a set of instructions to be followed by the reader. Similar instructions are found across Moore’s writings on perception (e.g. Moore 1953: 29-30). The supposed result of following this method is the identification of a truth about the nature of perceptual experience, namely that it involves a relation to sense-data. Moore’s philosophical argument involves asking the reader to engage in a certain kind of first-personal reflection, the result of which will be to disclose a philosophical truth about the nature of perception. And he presents this first-personal reflection as part of what is to be understood as involved in the method of analysis.

How should we understand the knowledge which results from this process? I noted that Moore’s dissertations make a distinction between analysis and analytic truths such that not all instances of analysis result in knowledge of analytic truths. Moore’s analysis of the proposition ‘This is a hand’ suggests one reason why this is so. For if one’s first-personal perspective is exploited in recognition of the fact that perception is a relation to sense-data, then we can see why Moore might distinguish it from analytic knowledge: it does not involve mere explication of that which is involved in the concept of perception. This is a way of understanding Moorean analysis which makes clear why Moore denied its results the status of analytic knowledge.

Does analysis, then, result in empirical knowledge? This is a more delicate question. It is important to Moore that the knowledge delivered by first-personal reflection is not the kind of knowledge which requires investigations
into the mechanics of the perceptual system. But a Kantian might claim that it is still the result of the deliverances of inner-sense and in this sense counts as empirical. This was one reaction of those working in early twentieth-century philosophy and psychology who took disputes about the knowledge delivered through first-personal reflection to motivate scientific study of the mechanisms of introspection and their reliability.  

An alternative is to think of first-personal reflection as delivering knowledge about the nature of perception in a way that precludes the characterisation of such knowledge as empirical. Husserl, writing at a similar time to Moore and on a set of related issues, endorsed the idea that phenomenological reflection could identify the essential features of experience. This reflection is supposed to be both distinct from and more fundamental than any empirical psychological investigations of experience. If Moore also thinks that the involvement of the first-personal perspective in philosophical theorising ensures that the knowledge gained is non-empirical, then we have an explanation for why he thinks his philosophical claims about the nature of perception are both non-analytic and yet in some sense a priori.

I won’t make the case that Moore endorses this claim here. But even if it can be defended, there are deep and difficult issues about whether the resulting view is coherent. If the thought is that first-personal reflection does not deliver empirical knowledge on grounds of enabling knowledge of natures, then there needs to be a reckoning with Saul Kripke’s identification of truths about essence which can be known through experience. If Husserl’s model of phenomenological reflection is a model for how such first-personal reflection works, then there is a question about whether it can be detached from the idealism about the objects of first-personal reflection which Husserl seems to

---

22 See (Spener 2018) for the details.
23 See especially (Husserl 1917).
24 These commonalities should serve as a reminder that the phenomenological and analytic traditions are but two stems of philosophy with a common (and known) root in Kant. See (Martin 2003) for discussion of Moore and Husserl and (Gomes 2017a) for discussion of the Kantian influence on early-analytic philosophy of perception.
25 See (Baldwin 2013; Martin 2003) for discussion. (Gomes 2017b) compares the methods of Moore and Strawson in more detail.
26 (Kripke 1980)
endorse in his later writings.\(^\text{27}\) And independent of both of these issues, one might still want some story about how it is that the use of first-personal reflection in philosophical theorising results in knowledge which is both non-analytic and a priori.\(^\text{28}\)

Still, without pretending to address these issues, we have enough to return to Strawson. For the suggestive thought we have extracted from Moore is that one might take the involvement of first-personal reflection in analysis to explain why its deliverances should be classified as both non-analytic and yet a priori. This does not yet make the category unmysterious. But it opens up a route for understanding Strawson. For if Strawson shared this view—that first-personal reflection can deliver non-analytic but a priori knowledge of philosophical truths—then we can make sense of his puzzling claims about the methodology of *The Bounds of Sense* so long as that methodology involves essential use of first-personal reflection. Does it?

\section*{5. Strawson (II)}

On the face of it, *The Bounds of Sense* does not contain any of the explicit instructions to attend to one’s own perspective which characterise the writings of Moore and Husserl. Here is how Strawson opens the book:

> It is possible to imagine kinds of world very different from the world as we know it. It is possible to describe types of experience very different from the experience we actually have. But not any purported and grammatically permissible description of a possible kind of experience would be a truly intelligible description. There are limits to what we can conceive of, or make intelligible to ourselves, as a possible general structure of experience. The investigation of these limits, the investigation of the set of ideas which forms the limiting framework of all our thought about the world and experience of the world, is, evidently, an important and interesting philosophical undertaking. (1 §)

There is nothing comparable here to Moore’s instructions for revealing sense-data, no precept to the reader to examine their own perspective, and no

\(^{27}\) See (Bell 1990, §III) for discussion.

\(^{28}\) See (Spener 2018) for discussion of the way in which this question shaped the development of early twentieth-century philosophy and psychology.
indication from Strawson that he takes the truths he identifies to be established directly through simple reflection on the character of one’s experience.

Nevertheless, once we look at the details of Strawson’s arguments, we can see a way in which one’s first-person perspective plays an essential role in establishing the conditions in his reconstructed Deduction. Consider the last of those claims: that possession of a conception of objectivity requires experience of objective things. How does Strawson establish this condition? He does not do so directly, drawing our attention to some manifest truth about the character of experience. Rather he argues for what he takes to be the contrapositive: that subjective experience is incompatible with the possession of a conception of objectivity (98-100). Sense-datum experience, Strawson says, consists of ‘impressions which neither require, nor permit of, being “united in the concept of an object” in the sense in which Kant understands this phrase’ (99). If subjects had nothing but this form of experience, they would not possess a conception of objectivity. So subjective experience entails lack of a conception of objectivity. Strawson takes this as equivalent to the claim that possession of a conception of objectivity requires objective experience.29

How does Strawson establish that purely subjective experience would lack a conception of objectivity? This is a counterfactual about what would be the case if perceptual experience lacked some feature which it actually has. A natural thought is that we determine its truth by engaging in imaginative reflection about how things would be were perceptual experience to be different. Imagine that you have a sense-datum experience. What would you be able to do in that scenario? Answering this question requires us to draw on the knowledge we possess of the character of our experience. We exploit that knowledge in determining the kinds of things which would be possible were experience to lack the character that it actually has. So in coming to know the counterfactual at the heart of Strawson’s argument, we have to exploit our first-person perspective on the world. The first-person perspective is thus

29 For some concerns about whether these claims are equivalent, see (Gomes 2016: 957-960) and (Gomes 2017b: 141-145).
central to an account of how we know the counterfactuals which make up Strawson’s reconstruction of the Deduction.  

Indeed, it is central to the arguments of Strawson and his students more widely. In *Individuals*, Strawson argues that (the perception of) space is a necessary condition on the reidentification of particular bodies. And he does this by imagining a sound world in which we do not perceive spatial properties and asking whether we would be able to reidentify particulars in such a scenario (Strawson 1959: 59f.). In his commentary on this chapter, Gareth Evans imagines a world of purely sensory properties and asks whether we could extract the idea of an objective property from such a world (1980: 98f.). John Campbell, in defending a relational conception of experience, asks whether our conception of a mind-independent world could be made available by the experience of a conscious image, before concluding that it could not (2002: 134-135). In each case, the main condition in some argument is supported by imaginative reflection on the kinds of things we would be able to do were experience otherwise. These exercises of imaginative reflection draw on our knowledge of the character of our experience. They exploit our first-person perspective on the world.

What are the implications for Strawson’s response to the metacritical challenge? We noted above that Strawson disavows that his claims are instances of analytic knowledge. And if the above account of Strawson’s methodology is along the right lines, then we have the shape of an explanation as to why this should be so. Both Moore and Strawson hold that analysis need not result in analytic truths. In the case of Moore, I suggested that this thought is underwritten by the way in which his use of analysis involves essential use of one’s first-personal perspective. The same, we can now see, holds true of Strawson: it is because knowledge of the character of experience is exploited in our coming to know the truths of various counterfactuals that the claims which make up Strawson’s arguments in the Deduction should not be characterised as analytic knowledge.

But Strawson also takes his claims to be non-empirical, in some broad sense. And one might worry that if they are supported by a process of imaginative

---

30 This interpretation of Strawson is defended more generally in (Gomes 2017b).
reflection, then they should count as empirical. Perhaps this is because such reflection involves empirical mechanisms of introspection, as many in early twentieth-century psychology alleged.\textsuperscript{31} Or perhaps this is because the only way such claims would fail to be empirical would be if some form of idealism were true, as some readers of Husserl have alleged.\textsuperscript{32} Or perhaps one might simply claim that the distinction between the a priori and the empirical loses significance at exactly this point, once one recognises the role that imagination plays in counterfactual reasoning more generally.\textsuperscript{33}

These are important challenges to the coherence of Strawson's view. But they are not challenges to its attribution. Rather, they demonstrate the difficulty in deciding whether the involvement of one's first-person perspective in philosophical theorising suffices to insulate philosophical reasoning from the methodologies of natural science. These issues were central to early twentieth-century debates in philosophy and psychology. They were less central, if present at all, in the debates occasioned by Strawson's use of transcendental arguments in\textit{Individuals} and\textit{The Bounds of Sense}. If what has been said here is along the right lines, that is a lacuna. Both Moore and Strawson want a philosophical method which is more than simply the unpacking of definitional truths but is yet distinct from the methodologies of natural science. My suggestion has been that if Moore and Strawson thought themselves to have such a method, it is because of the role that the first-person perspective plays in their philosophical theorising. And it is a deep and difficult question whether such appeal to the first-person perspective is enough to underwrite philosophical autonomy.

Let me draw these threads together. I have suggested that Strawson's criticism of the imaginary subject of transcendental psychology is best understood as an instance of metacritique, challenging Kant to provide grounds for the claims which make up his argument in the\textit{Transcendental Deduction}. Strawson's own reconstruction of the Deduction is open to such a challenge and he does not evade it by taking the claims which make up his own

\textsuperscript{31} See (Spener 2018) for discussion.
\textsuperscript{32} See (Zahavi 2007)
\textsuperscript{33} See, especially, (Williamson 2007, 2013). Williamson takes his account of the role imagination plays in our knowledge of modality to undermine the significance of the distinction between the a priori and the empirical.
argument to be analytic. Rather he puzzlingly commits to our possessing non-analytic but still a priori knowledge of their truth. I've suggested that this may be a result of the role that our first-personal perspective plays in establishing those claims. And although this does not make the view any less mysterious, it at least situates him within a recognisable tradition in twentieth-century philosophical theorising. For Strawson, we might say, it is the involvement of the first-person perspective in philosophical theorising which distinguishes it from both analytic explication and empirical science.

6. Kant (11)

This would be an appropriate note on which to end a paper for a volume on Strawson. But chiasmus demands that we end with Kant. For one way to read Kant's own route out of the metacritical trilemma is to hold that he too thinks we have a source of non-analytic a priori knowledge which can be appealed to in explaining the possibility of synthetic a priori knowledge. It is Kant after all who asks us to imagine the kinds of things that a being without spatial representation could do (A23-25/B38-39), to imagine the representation of a body without impenetrability, hardness, colour etc. (A20/B35), and to imagine deviance in the appearance of cinnabar (A100-1). And all of this in service of the identification of necessary conditions on our representation of objects. These processes look to involve the kind of imaginative reflection I have identified in Strawson. And one might take them to show that Kant allowed us to have synthetic a priori knowledge of the claims which underwrite the project of the Critique of Pure Reason.

Still, the metacritical challenge had force for Kant because it seemed that Kant was precluded from taking his claims about the structure of the cognitive faculty to be both synthetic and a priori on grounds of undermining his explanation of the possibility of synthetic a priori knowledge. So if the metacritical challenge is going to be answered in this way, we need a principled explanation for why some kinds of synthetic a priori knowledge do not need explanation and can thus be appealed to in the explanation of some others. Only so can the challenge be disarmed.

Strawson's version of the metacritical challenges focuses on the subject of transcendental psychology, that is, on Kant's claims about the structure of the
cognitive faculty and its role in legislating experience. And if it is these claims that Kant thinks are known both synthetically and a priori, then one natural way to distinguish between the problematic and unproblematic synthetic a priori claims—which is to say, those which need and those which can be appealed to in explanation—is in terms of those which concern the structure of the mind and those which do not.³⁴ Kant sometimes uses the term *reflection* to pick out that *state of mind in which we first prepare ourselves to find out the subjective conditions under which we can arrive at concepts* (A260/B316). And one answer to the metacritical challenge is to say that the structure of the cognitive mind is known, for Kant, through reflection. And that there is nothing problematic in appealing to such knowledge in an explanation of those synthetic a priori claims which cannot be established by reflection.³⁵

It requires work to see if this suggestion can bear weight. But if it does, then there is a pleasing convergence in the methodologies of Kant and Strawson. It is Strawson, of course who co-opts Kant into his methodology, classifying him with Aristotle in the Preface of *Individuals* as a purveyor of descriptive metaphysics (1959: 9). This has always been a source of bafflement. How could Kant, that transcendental idealist who thought that spatio-temporal objects are nothing but appearances (A79-80/B158-9), be a descriptive metaphysician? The suggestion sketched above suggests an answer. Kant must distinguish the grounds for his claims about the structure of the mind from the grounds for his claims about the structure of the world. Strawson is thus half-right: Kant has a *descriptive metaphysics of mind* but he combines it with a *revisionary metaphysics of nature*.³⁶

³⁴ See (Smit 1999; Marshall 2014) for versions of this response.
³⁵ See (Smit 1999; Westphal 2004: 12-32; Marshall 2014: 564-567; Merritt 2018) for discussion of the role reflection plays in Kant’s arguments and (Gomes 2017b) for wider discussion.
³⁶ Thanks, as always, to Andrew Stephenson for discussion and comments and to the Editors of the volume for their comments.
References

Works by Kant:

References to the *Critique of Pure Reason* follow the standard A/B pagination. All other references are to the volume and page number of *Kants Gesammelte Schriften* (Berlin: De Gruyter and predecessors. 1900—). Translations have been from the following works in the *Cambridge Edition of the Works of Immanuel Kant* (Cambridge: Cambridge University Press):


Strawson’s quotations from the *Critique of Pure Reason* are taken from the translation by Norman Kemp Smith (London: Macmillan, 1933).

Works by other authors:


