

ARTICLE

Kant and the transparency of the mind

Alexandra M. Newton 

Department of Philosophy, University of Illinois at Urbana-Champaign, Champaign-Urbana, IL, USA

ABSTRACT

It has become standard to treat Kant's characterization of pure apperception as involving the claim that questions about what I think are transparent to questions about the world. By contrast, empirical apperception is thought to be non-transparent, since it involves a kind of inner observation of my mental states. I propose a reading that reverses this: pure apperception is non-transparent, because conscious only of itself, whereas empirical apperception is transparent to the world. The reading I offer, unlike the standard one, can accommodate Kant's claim that the I of pure apperception is the same as the I of empirical apperception.

ARTICLE HISTORY Received 18 July 2018; Accepted 2 January 2019

KEYWORDS Self-consciousness; transparency; apperception; Kant; Sartre; Evans

When I am asked whether I think there will be a third world war (when asked about my thoughts), I do not answer this question by looking inwards, at the contents of my mind, but by directing my attention outwards, at current events that might portend the advent of another world war. This phenomenon, which has come to be known as the *transparency* of questions about the mind (about what I think or judge) to questions about the world (about what is the case), is often thought to be central to Kant's reflections on self-consciousness in theoretical judgment. When he introduces transparency, Gareth Evans, for instance, says 'I believe we may have here an interpretation of Kant's remark about the transcendental "I think"' (Evans 1982, 228). Kant says that the 'I think' that expresses transcendental apperception (or self-consciousness) is a 'wholly empty representation' (A345-6/B404). It is empty, according to transparency views, because we see *through* ourselves to the world; we do not make ourselves an object of attention, but are conscious of our thoughts and judgments by directing our attention outwards, at the world we think about. These claims are usually taken to be restricted to what Kant calls 'transcendental' or 'pure apperception' (A107, B132). 'Empirical apperception' or inner sense, by contrast,

CONTACT Alexandra M. Newton  amnewton@illinois.edu  Department of Philosophy, University of Illinois at Urbana-Champaign, IL, USA

© 2019 Canadian Journal of Philosophy

involves directing my attention inwards, and passively observing an inner, temporally ordered stream of consciousness.¹

Kant famously maintains that the I of pure apperception is 'identical' with the I of empirical apperception 'as the same subject' (B155). There are not two I's here, but one. As he notes in the *Anthropology*, this gives rise to a puzzle: the "I" appears to us to be double (which would be contradictory): 1) the "I" as subject of thinking (in logic), which means pure apperception (the merely reflecting "I"), and of which there is nothing more to say except that it is a very simple idea; 2) the "I" as object of perception, therefore of inner sense, which contains a manifold of determinations that make an inner experience possible' (7:134n.). As transparency accounts of pure apperception understand the puzzle, it is this: how can the I that always recedes behind consciousness be the same as the empirical I that appears as an object of consciousness? To put it in terms used by Wittgenstein, how can the I that is the 'limit of the world' also be an I that 'belongs to the world' (Wittgenstein 2001 *Tractatus*, 5.632)? Or in Sartre's terms, how is the I that is nothing (i.e. is not posited as an object of consciousness) the same as the I that is something (i.e. is posited as an object of consciousness)? As we shall see in the first section of this paper, transparency readings of Kantian pure apperception have no way of answering these questions.

In the second section I will argue that the identity of the pure and empirical I's can be made intelligible only if we revise the accounts of pure and empirical apperception that are presupposed by transparency views and standardly assumed in the literature on Kant. First, with regard to pure apperception, we must call into question the subject-object (or force-content) distinction that is dogmatically assumed by transparency views. I will argue that the object is not originally distinct from the act of thinking in transcendental apperception. So thinking does not go outside of itself in thinking its object. Second, I will question the view that empirical apperception is a kind of inner observation of my mental states. In a reversal of the standard reading, it will emerge that transparency plays no role in transcendental apperception but is central to Kant's account of empirical apperception. The *empirical self* is fleeting and impossible to grasp as a persisting object; we are conscious of our representations in inner sense only by directing our attention outwards, at persisting objects of outer sense. Rather than thinking of the empirical self as an object of reference, a particular given to consciousness as something particular, I will argue that the empirical I is the universal I of pure apperception in its act of self-particularization.² This will enable us to appreciate how the I of pure apperception is non-accidentally the same as the I of empirical apperception.

1. Transparency accounts of pure apperception

There are two broad approaches to the transparency of pure apperception – one of which I will call ‘Evansian’ after Gareth Evans, the other ‘Sartrean’. The Evansian approach assumes that there are particular beings in the objective, spatio-temporal order that are able to refer to objects, including to themselves. We can ask *how it is* that they are able to refer to themselves when they self-ascribe mental states in such thoughts as <I think p >.³ Transparency characterizes the peculiar ‘procedure’ that they employ in such self-ascriptions, distinct from the methods they use to become aware of properties ascribed to objects other than themselves (or to objects that just happen to be themselves) (Evans 1982, 225).⁴ On the Sartrean approach, by contrast, the transparent I does not refer to itself or identify itself with a being in the world, but distinguishes itself from all beings in the world. Transparency thus is not a method used in self-ascriptions of mental states, but characterizes self-consciousness *without* self-reference. Whereas on the Evansian approach, the transparent I presupposes an empirical I that is in the world, the Sartrean approach reverses this: the empirical ego, as an object of what he calls ‘thetic consciousness’, presupposes the transparent I (the I of non-thetic self-consciousness).

It is not my intention in this paper to settle exegetical questions about Evans’s or Sartre’s own accounts of self-consciousness. My main purpose is to discuss their appeals to Kant, and to assess whether their transparency approaches to Kantian pure apperception are compatible with the sameness of the transcendental and empirical I’s.

The Evansian approach, which has come to dominate much of the literature on Kant, assumes a distinction between two standpoints: one, external, from which I think of an objective world (Evans calls this the ‘fundamental level of thought’ (Evans 1982, 105f., 152, 210)); the other, internal, from which a being within this world thinks of herself first-personally, or from a first-person perspective.⁵

Evans notes that in thinking about myself *from within*, I ascribe thoughts to myself in a manner very different from the way that I ascribe properties to outer objects. It is not by observing myself thinking <p>, or by using criteria for identifying myself as the one who thinks <p>, that I know I think p. It makes no sense to be aware of the thought <p> without being aware that I am the one thinking it; I don’t have to identify myself as the thinker of this thought in order to know that it is mine. I am aware of <p> as mine by directing my attention solely at p itself. The thinker of <p> can thus be said to be transparent to the world: for in consciousness of thinking p, it is not conscious of itself as an object of thought at all, but recedes *behind* what it thinks. It is in this context that Evans says ‘I believe we may have here an interpretation of Kant’s remark about the transcendental “I think”’ (Evans

1982, 228). The transparent or transcendental I is not an object of reference and so does not enter the content of my thoughts, but is, as Evans says, 'purely formal' (Evans 1982, 226).

However, Evans thinks that this formal self-consciousness is not sufficient to 'constitute a full understanding of the content of the judgment "I believe that p"', since the content of that judgment involves the use of I as a referring expression (Evans 1982, 226). To think that transparency is sufficient to secure the referential significance of 'I' would be to make the Cartesian mistake of inferring from thinking alone to an object or substance that thinks. Evans follows Strawson in reading Kant's core insight (against the rationalists) to be that self-conscious thought is inseparable from a conception of oneself as an embodied being located within an objective world: transcendental apperception is inseparable from empirical apperception or inner sense. For if we could not situate ourselves within an objective world-order that is there independently of our perspective from within it, we would not have a capacity to refer to ourselves in self-ascriptions of thoughts or beliefs.⁶ It is in the self-locating thought of empirical apperception, through which I become conscious of my spatio-temporal position, that Evans thinks I identify myself with a being in the world. Empirical apperception thus establishes self-reference.⁷

Notice that the transparency method for ascribing particular thoughts to myself will only work if the I that thinks <p> is conscious of itself as the same as the I in the self-ascription <I think p > . This is often overlooked, as though the puzzle of transparency that Evans's account raises could be that of how I can go from a mind-independent fact about the world, <p>, to a fact about my mind, <I think p > .⁸ The real puzzle does not concern the relation between these thought-contents (whether inferential or not), but rather between the I that disappears – the *transcendental* I that thinks <p> – and the I that appears or that becomes the object of self-ascription – the *empirical* I of <I think p > . How am I conscious that the I of <I think p> is the *same* as the I that thinks <p>? Evans tries to avoid this difficulty by assuming an empirical I and by thinking of the transcendental, transparent I as a merely abstract aspect of it.⁹ The transcendental I of <I think>, he says, is embedded in, or conceivable only against the 'background' of, the use of I in self-locating thought (Evans 1982, 226). For instance, in thinking <I think p>, I am ascribing the thought <p> to myself, as an empirical being in the world, distinct from others. The point of the transparency method is that I can do this in abstraction from any consideration of that which distinguishes my thinking <p> from your thinking <p > . I can do it simply by focusing my attention on p. It is thus by abstracting from the empirical foundation of I-thoughts that I form the idea of an entirely empty or formal I, a 'consciousness in general' that is *impersonal* and universal.

But once we take this abstractive approach, it emerges that the transcendental I is not transparent as the account purports it to be. To borrow Sartre's words from a different context, the empirical I cuts through the transcendental I like an 'opaque blade' (Sartre TE, 40).¹⁰ For the subject of thinking – properly understood – now just is the empirical subject, and its acts of thinking, which were operative in thinking <p>, are psychological acts ascribed to this individual subject as its properties. Thus, these mental acts were already there, cutting through my consciousness of p, even though I was abstracting away from them. The subject therefore always thinks the outer world through its subjective, limited perspective, from a place within it. It thinks *what is* from the perspective of how things *seem to be* to it. One cannot think of the objective world (what is) as an abstraction out of the subjective world (what seems to be). If the objective world is thought of as an abstraction, as, for instance, what all 'seemings' share in common, it is not truly objective. Likewise, if the I that thinks the world is an abstraction out of the empirical I, it is not truly transparent to the objective world. Evans wants to be able to say that the objective world-order is already there, independent of perspectives of subjects within it. But by placing the subject within an objective context, he has placed it in a jail from which the subject cannot free itself by simply trying to 'abstract away' from the fetters of its confinement. The subject is not in a position, from within the confinement of its subjective perspective, to distinguish how things *seem to be* to it from how they objectively *are*.

In readings of Kant, this approach to transcendental apperception would have the disastrous consequence that transcendental idealism is a subjective idealism. The first *Critique* adopts the first-person standpoint of the thinker in determining ways in which it thinks about or gains cognitive access to the world. But on this reading, the transcendental structure of the world would be the formal structures discovered from within an externally limited, subjective perspective of a being in the world. One might say that subjects within the world have objective knowledge of the world around themselves – but one cannot *entitle* oneself to that claim if one has already confined the subject to a perspective from within it.¹¹

The Sartrean approach is more Kantian in that it does not dogmatically assume a distinction between external and internal viewpoints, or between the objective world-order and the first-person perspective of a being within that world-order, but allows philosophical reflection to begin from within a standpoint that transcends this distinction and first makes it possible.¹² Sartre notices that thought about an objective world-order is inseparable from self-consciousness in thinking it. All consciousness is consciousness of something other than itself (it is characterized by *intentionality*); in Sartre's words, it is *thetic*, for it posits an object beyond itself. But it is also (non-thetically) conscious of itself *as of* something other, and thus of itself as distinct from the object of consciousness.¹³ But in non-thetic consciousness

of itself it is not conscious of itself as having any given characteristics, or of being limited in any way. Rather, consciousness is conscious of itself as *nothing* in contrast to the '*something*' it thematically posits or is conscious of: consciousness is, as Sartre puts it, 'all-lightness, all transparency' (*toute légèreté, toute translucidité*) (Sartre TE 1960, 42; altered translation; BN, 1956, 78).¹⁴ So self-consciousness is completely transparent to the world because to be self-conscious is to be conscious of escaping ourselves (going outside ourselves, transcending ourselves) in being intentionally directed at outer objects or at states of affairs.¹⁵ Indeed, transparency is a poor metaphor for characterizing the subject on this view, insofar as it leaves in place the idea of a medium through which we look outwards, at the world.¹⁶ When we go outside, there is nothing, not even a medium, left inside (there is no opaque blade that cuts through its consciousness of the world).¹⁷ Everything, on the Sartrean view, including mental states, is external to the non-thetic 'I', because everything *transcends* our consciousness of it: 'everything is finally outside, everything, even ourselves. Outside, in the world, among others' (Sartre 1970, 5).

The I of non-thetic self-consciousness, according to the Sartrean approach, is not a thing (spatio-temporal object) that I am conscious of in a special way. Non-thetic self-consciousness is not a peculiar kind of thought *about* myself (with a special Fregean sense), but is the entirely empty form of all thinking (indeed, of all consciousness) in general. Insofar, then, as the 'I think' is an expression of non-thetic self-consciousness, <I think p> means just the same as <p>: the <I think> does not add any additional content to that which I think, namely, to <p> – the <I think> expresses *nothing*. Thus, the non-thetic <I think p> is not a self-ascription in Evans's sense, but accompanies any thought in general (not just thoughts about myself).¹⁸ Contrary to the Evansian approach, Sartre maintains that 'consciousness ... can be limited only by itself' (Sartre TE 1960, 39). So the limits of the empirical ego are *inside* consciousness; only within consciousness can we draw a distinction among the objects of thought, between an ego that is a being in the world and others outside it, or between the way things appear to me, as a particular in the world, and the way they are outside that perspective.¹⁹

In the *Transcendence of the Ego* Sartre suggests that this nothingness of consciousness in relation to its object is the proper way to understand Kant's thesis that the 'I think' is an empty representation. Insofar as general and transcendental logic are concerned with mere empty forms of thinking and judging, they must be concerned with the merely 'logical conditions' that are 'necessary for the existence of an empirical consciousness', i.e. conditions for intentional relations to objects (Sartre, TE 33). These are not psychological conditions or given characteristics of the subject; the 'transcendental field', as Sartre says, is 'impersonal; or if you like, pre-personal, without an I' (Sartre, TE 36). On Sartre's reading, they can be understood as

the formal structures of intentionality, i.e. ways in which consciousness (or thinking) is directed at the world (or being).

Notice that Sartre does not face the difficulty of explaining how the I that thinks is the same as the I that perceives or the I that runs, since all acts of consciousness – whether in thinking, perceiving, or (practical) intention, are intentional acts, and thus have the same ‘nothingness’ – the same *non-thetic I* – at their core. But on Sartre’s view the non-thetic I cannot become an object of experience: the I that is *nothing* is not the same as the posited ego that appears and is *something*. Sartre says that the latter is a transcendent object just like any other object of consciousness, and so is distinct from consciousness of it: the transparent I can always say ‘I am not what I am’ (Sartre BN, 260). Indeed, Sartre suggests that this ability to step back from, or to negate, the object-ego, constitutes our freedom. The sense in which I can identify myself with the object-ego, as when Sartre says ‘I am what I am not’, is very different from the sense in which I am *not* identical with it (‘I am not what I am’, *ibid.*). In *Being and Nothingness*, he suggests that I identify myself with an object-ego in the way that the craftsman identifies himself with the tool in his employment of it. The tool is not given to him as an alien object, but is the instrument of his engagement with the world (Sartre, BN 303ff.).²⁰ (This is the kind of self-identification that Korsgaard appeals to when she says that I can identify myself with my vacuum cleaner (Korsgaard 2009, 37)). In the *Transcendence of the Ego*, Sartre puts this by saying that I can become ‘intimate’ or familiar with an ego-object as my own (Sartre TE, 86). But notice that this still leaves us with two I’s: one, which is truly a subject-I and transparent to the world, and another, which provides a familiar dwelling for it within the world. Intimacy is not sameness.

As a reading of Kant, the Sartrean approach thus leaves us without resources to interpret Kant’s claim that there is no ‘doubling’ of I’s. According to Kant, there are not two I’s here, but one. It is *I* who am affected in sensory perception, not merely a body that I occupy. Indeed, my actions would be unintelligible as *mine* if I could not say that I *am* this individual, particular self that exists in the world, but could only say that I *am intimate with* this individual. For surely it is *I*, not a person I occupy, who acts in the world.²¹

We have seen with Sartre, against Evans, that the transcendental I cannot be a mere abstraction out of the empirical I and be fully transparent; on the contrary, the empirical I only comes into view for a transcendental I, and thus presupposes the latter. But the Sartrean approach makes it seem that any attempt to *identify* the transcendental and empirical I’s as the *same* I would be a confusion. For, on this approach, the transcendental I is the empty subject of an act of intentionality, which is universal and formal, while the empirical I is one of the many particular objects that it can become conscious of. In what sense of *identity*, then, can Kant claim without confusion that they are identical? I shall return to this question in [Section 3](#).

2. Kant on transcendental apperception

First, we would be well advised to turn to Kant's own reflections on transcendental apperception in the first *Critique*, since the above accounts may have been led astray by assuming that transparency plays any role in it. My remarks here will be brief, as my main goal is to articulate the relation between transcendental and empirical apperception (in the next section). We saw that transparency accounts of pure apperception characterize thinking primarily through *intentionality*: to think is to be conscious of something distinct from the thinking of it. I am conscious that I think <p> by going outside myself and focusing my attention on p. Since, by thus going outside, I leave nothing inside myself, the I could be said to be transparent to the world.

However, Kant's fundamental conception of thought doesn't involve thinking of thought as an act of intentionality. This is especially evident in his general logic, where thinking is understood as a (formal) act or function of synthesis, in abstraction from the relation it bears to any object. But even in transcendental or material logic (as expounded in the B-edition of the first *Critique*), which I will focus on here, Kant does not originally think of the relation that thought bears to its object as one of intentionality. It is not in the *pure* self-consciousness of judging, but in the *empirical* awareness of a temporally extended act of judging, that Kant discusses judgment's intentional relation to an object that is 'outside' of it. One should not confuse the relation that judgment originally bears to an object at the transcendental level (of pure apperception) with the intentional relation that representations in inner sense bear to the objects of outer sense at the empirical level, or so I shall argue. In the former relation, the act of judging does not go 'outside' itself, but remains entirely immanent to itself.

In the B-edition Transcendental Deduction, Kant glosses the 'relation of representations to an object' as their 'objective validity' (B137). Objective validity is what is expressed by the 'copula **is**' in a judgment (B141f.). A judgment thus 'relates to an object' insofar as it is objectively valid. Kant later clarifies that if a judgment is objectively valid, we can infer that it is true (A125, A788/B816, cf. A131/B170, A202/B247). For a judgment is objectively valid only if it accords with the laws of the understanding (both logical and transcendental), and no lawful exercise of the understanding (as a capacity for knowledge) can be false (A293-4/B350). The relation to the object therefore must include 'the agreement of cognition with its object', which is Kant's nominal definition of truth (A58/B82).²²

When a judgment is objectively valid, the representations within it are related to 'the original apperception and its **necessary unity**' (B142); i.e. it is not accidental or haphazard that I combine representations in this manner, rather I combine them from a consciousness that they belong together necessarily, not just for me, but for any subject of judgment. Judgment

thus involves a consciousness of itself as not just true, but non-accidentally true, i.e., as knowledge. When Kant says that false cognition 'does not agree with the object *to which it is related*' (A58/B83, my emphasis), this does not mean that a cognition relates to an object (or is objectively valid) *insofar* as it is false, but that it relates to an object insofar as it is true: 'in every erroneous judgment there must always lie something true' (JL 9:54). For instance, the false judgment that the table is brown may contain the truth that the object is a table, and thus may still contain relation to an object (objective validity).²³

Some interpreters have rejected this interpretation of 'objective validity', since Kant says that all judgments have objective validity (B142); it would thus have the absurd consequence that all judgments are true.²⁴ However, it should come as no surprise that the analytic of transcendental logic ('analytic of truth') considers all judgments to be true, if we read it as taking a first-person standpoint on judgment. Within the self-consciousness of an act of judging, as an act of the capacity to know and thus to judge truly, I cannot think of my judgment as false or even as possibly false. For to judge is to make a claim about the way things are; it is to think of one's judgment as true.²⁵ I cannot judge <S is P> and, in that same act, think that <S is P is false> or even <S is P may be false>. It is only when I step back and consider *whether* I should make a judgment, or when I reflect on how I have judged in the *past*, and thus take an external perspective on judging, that I can think of a judgment as possibly false.²⁶

Indeed, the reason Kant takes this first-person standpoint on judgment is that there is no external or *sideways-on* perspective on the truth of cognition at all: this is what Kant suggests by identifying the objective validity of a cognition with 'objective unity of the *apperception* of the concepts contained therein' (my emphasis; B140f.). Agreement with the object is not a relation that acts of judging bear to an external object, a relation that can somehow be viewed from outside of apperceptive acts of judging: for 'I can compare the object with my cognition [...] only by cognizing it' (JL 9:50; cf. A57-58/B82-83). My judgments of the object can only be corrected by other *judgments* of it. The relation of *agreement* with the object is thus *internal* to the relation of acts of judging (cognizing) amongst themselves, and so is internal to the apperception or self-consciousness of such acts.²⁷

There are two ways to read this internality: either truth is internal to apperceptive acts of judging because the primary bearers of truth are the *contents* of these acts, or because the primary bearers of truth are the *acts* of judging themselves. Along the lines of the first interpretation, John McDowell has suggested that for Kant, as for Wittgenstein, truth is internal to the act of judging because when I judge truly, my mind does not stop anywhere short of the world. What I judge – the content of my judgment – is identical with what is the case. Thus, there is 'no gap between thought, as

such, and the world', understood as 'everything that is the case', and hence no need for an external relation of correspondence between what is inside of my judgment (its content) and an object external to it (McDowell 1994, 27). Rather, there is an identity of what I judge and what is the case in the world.²⁸ Although this eliminates the 'sideways-on' conception of truth as an external relation between contents inside the mind and objects outside of it, it still retains a distinction between the *act* of judging (force) and the *content* judged. The truth-bearers, on McDowell's reading, are the *contents* of judging, not the acts of judging them. Acts of judging can be said to be true in the sense of correct only insofar as the contents of judgments are true, i.e., are identical with states of affairs. This is in line with transparency views of transcendental apperception, since it is by focusing on something distinct from acts of the mind – namely on their true contents – that I become conscious of these acts as my own.

However, there is scant textual evidence that Kant, like his Fregean successors, distinguishes the act of judging from the content judged.²⁹ Moreover, when Kant says that objective validity is expressed by the 'copula **is**' in a judgment, this means that it is expressed by an *act* of judging (B142). For unlike the Fregean conception of judging as an act external to the content judged, Kant views the act of judging as an act of *synthesis* of representations *internal* to what is judged – precisely that act of synthesis that is logically expressed by the copula 'is'. The primary truth-bearers are acts of judging (of synthesis), not act-independent contents of judgment, since truth in the primary sense is objective validity, and it is the *act* of judging (internal to what is judged) that is true in the sense of valid or correct.

To be conscious of the truth of my judgment thus is not to direct one's mind towards an act-independent fact in the world, rather it is to be conscious of the act of combining representations in one necessary unity of consciousness, one judgment. A judgment <S is P> is true when the predicate <P> agrees with the object thought through the subject-concept <S> – i.e., when there is a necessary agreement among the representations combined within a judgment, in accordance with the logical and transcendental principles of possible experience. Agreement with the object is thus internal to agreement of the representations in a judgment amongst themselves, in the act of their synthesis.

Thus far we have focused on Kant's claim, from the first half of the B-deduction, that cognition's 'relation to an object' consists in its objective validity. Contrary to interpretations that read 'objective validity' as intentionality or objective purport, I have argued that objective validity is the non-accidental truth expressed by the copula 'is' in judgment, and that consciousness of being ('is') in the sense of being true (veritative being) and consciousness of thinking (or judging, synthesis) are the same.³⁰ Such acts of thinking (as objectively valid acts of synthesis) are only possible under conditions of a manifold given from elsewhere, but this does not entail that consciousness of the truth of the

thought is consciousness of something given from elsewhere. To be conscious of truth is not to reach out beyond the act of thinking or judging to an independent object that it is true of, rather truth (objective validity) is internal to the act itself (i.e. to the copula). Thus, to be conscious that my cognition 'relates to an object' in the sense of being 'objectively valid' just is to be self-conscious in the act of judging.

Now, the I of <I think> in such acts is formal or empty, but not because it is nothing *in contrast* with *something* (the truths) thought about; rather, it is empty because it is consciousness of truth in general (or of an 'object in general'), which is not actually, but is only potentially, all particular truths (about particular objects).³¹ The formal concept <I think> is the logical concept of the functions of thinking in judging (functions of logical synthesis) that are the bearers of truth.³² To relate this concept to objects is not to go outside it, but to fill it with transcendental and empirical content.³³ For contentful knowledge of objects does not require directing our attention away from logical functions and towards an external object, but becoming conscious of these same logical functions in their role of determining a sensible manifold, i.e. in their role of making particular, empirical truths (experience), and the objects of these truths (objects of experience), possible (KrV A158/B197).

Moreover, consciousness of the objective unity of various representations through <I think> is an impersonal, universal, and merely formal act of self-consciousness (see Engstrom 2013). It is not, as such, consciousness of a particular agent or thing that performs an act of combining, as this would restrict the truth (in the above sense) to what holds merely *for me*. As Kant emphasizes in the *Paralogisms* chapter, the <I think> expresses the formal logical unity of consciousness of subject and predicate concepts, i.e., the copula unifying the thought <S is P> – it does not add a substance or underlying subject to the thought itself (B406f.). The thought <S is P> and the thought <I think S is P> thus are the same thought; the latter does not bring attention to myself as an individual distinct from others. The <I think> merely expresses the logical, objective unity of the thought (objective synthetic unity of apperception), which is common to, or identical in, the elements in the thought, both in me, and in any other rational thinker (this sameness or identity Kant calls an 'analytic unity' of apperception). For I am conscious, in judging, that the judgment holds not just for me, but for any rational subject.³⁴ The logical or transcendental 'I' thus is not an expression used to refer to a particular individual, but is the common 'consciousness in general' that is shared by all possible (identical) acts of knowing, and thus by all truths (B143).

Although transparency views are right to emphasize that I am conscious of my judgments in transcendental apperception in an impersonal manner, merely through consciousness of what is true, this does not mean that I am conscious of something outside of the act of judging

itself – an act-independent content, for instance, or a mind-independent world-fact. Indeed, *nothing* at all – nothing, at least, that is anything ‘for me’ – is outside the transcendental I, understood as the ‘original-synthetic unity’ of all my representations in one consciousness (B132). The transcendental I is conscious only of *itself* and its own activity, and thus is non-transparent to any world-facts distinct from or external to itself and its activity.³⁵

3. Empirical apperception

What the above elucidation of Kant’s account shares with transparency views of self-consciousness is that in transcendental apperception of my thoughts and judgments, I am not conscious of any given character of myself, as the individual subject of thought.³⁶ I am conscious of thinking <S is P> just *in* thinking S is P, without at all thinking about myself as the individual thinking it. Indeed, I can also become conscious of what *others* think (nameless others) by abstracting entirely from the differences between them and myself, and focusing only on what I understand them to say. The experience of disagreement or error may direct our attention at these empirical differences, but within a setting where we are jointly determining what to think and we refrain from errors, it does not matter who says what: reason (consciousness in general) is determining what it thinks. It is not the transcendental subject, but empirical subjects – both myself and others – that are transparent, since reason (the transcendental I) ‘sees through’ them to itself alone in becoming conscious of what it thinks.³⁷

Contrary to logical and transcendental apperception, empirical apperception introduces an ‘inner’ realm of subjectivity distinct from the ‘outer’ realm of objects given to me in space (the form of ‘outer sense’). For Kant says that inner sense provides us with an ‘intuition of our self and our inner state’ (A33/B49; cf. A22/B37), while outer sense provides intuitions of outer objects. Whereas all being is immanent to the transcendental I, inner sense distinguishes my thoughts and perceptions, *my* being, from the being both of outer objects and of other subjects. As Plato and many classical philosophers noticed, the further we descend into our particular, empirical selves, the more distanced we become from what is. We should not cherish the false hope that we will establish contact with what truly is through our senses: truth is to be found ‘only in judgments’ (A293, B350).

Since we have identified being (what is) or truth with what is judgeable, we are now in a position to understand the division, within the concept of being (the object in general), between my being and the being of outer objects, as a division internal to judgment. Like all divisions in thought, this division is only possible under sensible conditions of a manifold given from elsewhere. But it is nevertheless an act of dividing (separating), an act of thinking; hence the

members of the division are not external to that act. Kant calls the empirical I an 'object' of experience because it is an object of thought (empirical judgment), like any other. Thought does not encounter particular objects of thought and the distinctions among them but is itself the source of these distinctions, insofar as it thinks not only the objects but also their differences. That is, the <I think> *specifies itself* into the thoughts <I think x> and <I think y > . As an object of thought, I am contained in this way under the universal, formal concept <I think>, just like any other object of thought.

In what sense, then, is the transcendental I identical only with *me*? How is my empirical I, which is *not* identical with the being of the outer object which it thinks, identical with the transcendental I, which *is* identical with the being of what it thinks? To understand this we must distinguish two senses of identity or sameness. First, there is the sameness of the I with anything it thinks. Just as anything that falls under a concept F must be 'homogeneous with' or at least partially the *same* as that concept F (since 'the concept must contain that which is represented in the object'), anything that is an object of <I think> – i.e., anything that is an object of thought at all – must contain the <I think> within it and thus be the same as it (A137/B176). Kant does not call this sameness with the <I think> *homogeneity*, but instead calls it an 'analytic unity of apperception', because the <I think> is not a *genus* common to a determinate sphere of objects, distinct from other *genera*. Rather, as we have seen, the <I think> is indeterminate or without limit in its extension: there is nothing outside of its realm (indeed, even a 'table of nothing' is *inside* its realm – cf. A292/B348). My empirical I, like all other empirical I's and all other objects of thought, is in this sense the *same* as the transcendental I.³⁸

But in the second sense of sameness or identity, only my empirical I is the same as the transcendental I. The key to understanding *this* identity is to see that my empirical I is not just *thought* by the transcendental I, but is also self-consciously *affected* by it. The empirical I becomes an object of thought through an act of self-affection – which is how the transcendental I situates *itself* within its own thought of an objective world. Whereas outer objects are given to me from elsewhere (in outer sense), only I am capable of being given to myself by myself (in inner sense). Rather than encountering myself as an already given particular existing in the world, I particularize myself, and, in inner sense, am able to become conscious of myself as an appearance in the world, in relation to others.

This approach can help us understand why Kant argues, in the Refutation of Idealism, that I do not have direct access to myself as an inner object that I encounter, but that I am conscious of myself in inner sense only through my active engagement with objects of outer experience. Very briefly, the argument proceeds as follows. I am conscious of my own representations only in their temporally determined relations to one another. But according

to the First Analogy, 'all determination of time presupposes something permanent in perception' – that is, it presupposes a substance (B275). This abiding substance cannot be 'something in me', since Kant shares Hume's insight (against the rationalists) that we do not experience ourselves as permanent in time; there is no intuition of ourselves as substances – everything in inner sense is 'in constant flux' (B291). Hence, the permanent substance underlying inner time-determinations must be an outer appearance: 'the determination of my existence in time is possible only through the existence of actual things which I perceive outside of me' (B275-6). I am temporally conscious of my own representations only by directing my attention outwards, at the objects of outer sense: 'The consciousness of my own existence is at the same time an immediate consciousness of the existence of other things outside me' (B276).

Thus, although inner sense is a kind of self-affection, it does not involve being affected by myself *qua* outer object (*qua* something *other*).³⁹ The objects of outer perception must be given to me, i.e. I must receive sensations or impressions from them by being affected by them, and there is no way to know that the object that affects me is the same as the subject affected (although the critical philosopher leaves this sameness an open possibility). As Wittgenstein suggests, although it makes sense to say, after a car crash: 'Here is a limb, but is it mine?', it does not make sense to say: 'These are representations, but are they mine?' (Wittgenstein 1958, 67f.). My access to my inner states is not like access to outer things, or to a body that just happens to be mine, but is first-personal self-awareness: I am aware of my representations *qua* non-accidentally *mine* – for instance, I'm aware of temporally extended acts of thinking as non-accidentally the same as the non-temporal, logical act of thinking.

Kant can acknowledge this first-personal character of empirical apperception because whereas sensations make up the materials of outer sense, 'the representations of outer sense make up the proper material with which we occupy our mind' (B67). That is, there aren't any special materials of inner sense. It's not that first, we are affected by the object in representations of outer sense, and then, in a separate act, we are affected again, now by the representations of these objects. In addition to outer sensations, there aren't any inner sensations.⁴⁰ I don't need to receive impressions from my inner states (perceptions, thoughts, etc.) in order to be aware of them as such. I can become conscious of intuiting and perceiving simply *by* intuiting and perceiving. In the case of thought, I am empirically aware of my thoughts by thinking them, not by stepping back from them and observing myself think them.⁴¹ And in perceiving a line, I am aware of perceiving it by 'drawing it in thought', not by stepping back and observing myself as I draw it.⁴² It's just that these activities through which I am empirically self-aware are

temporally extended, which is on Kant's view the same as saying that they're acts of self-affection.⁴³

In order to make sense of temporally extended acts of the mind ('motion [*Bewegung*]' of the mind, *kinesis*: B155), Kant thinks that we need to think of these acts as performed under the condition of an object affecting me (or having affected me) in outer sense. Temporal relations cannot arise *ex nihilo*, or from logical relations alone; that is, time is not a form of mere *spontaneity*, but of *receptivity*: temporal intuition rests on affection by objects outside of me. Therefore, I am conscious of my inner states in inner sense only by directing my attention outwards, at objects of outer sense.

We are now in a position to ask whether the transparency of *empirical* apperception is Evansian or Sartrean. That is, in directing my attention at objects of outer sense, am I a particular spatio-temporal perspective on the object, distinct from other perspectives (Evansian), or is my empirical I entirely empty – am I nothing at all, not even a particular perspective alongside others (Sartrean)? Sartre says 'all consciousness is positional in that it transcends itself in order to reach an object, and it exhausts itself in this same positing. All that there is of intention in my actual consciousness is directed toward the outside' (Sartre BN 1956, 11). This seems to be precisely the case for temporal consciousness: it is always escaping itself, constantly going from one moment to the next, and from one object to the next. I can say of myself (originally) only *that* I am, but not that I am anything in particular (B157n.).

On a Sartrean approach to Kantian inner sense, empirical self-awareness would be an awareness of its own 'emptiness'. But contrary to Sartre, we can now characterize this emptiness (or nothingness) as the emptiness of the *empirical* I and distinguish it from the emptiness of the *transcendental* I. Whereas the transcendental I is the emptiness of 'consciousness in general' that remains within itself, even in its acts of self-division and self-specification, the empirical I is the emptiness of temporal consciousness, which is constantly escaping itself. In transcendental apperception, I am not conscious of myself as distinct from others, or even from the objects of thought, whereas in empirical apperception I constantly distinguish myself (as existence) from anything that I am conscious of. Whereas the 'nothing' of transcendental apperception is the emptiness of universality, since the transcendental I can be shared by any thinker, the 'nothing' of empirical apperception is the emptiness of radical singularity, since the empirical I cannot be shared by any other thinker.⁴⁴ Transcendental apperception is empty as the capacity to (spontaneously) *determine* any object, whereas empirical apperception is the emptiness of the passive capacity *to be determined* by any such acts of determination.

The Evansian approach to the transparency of empirical apperception, by contrast, would identify my empirical I with a particular, determined spatio-temporal perspective in the world. In an influential paper, Markos Valaris

adopts this approach in his interpretation of Kantian inner sense: it provides us, he argues, with an awareness that 'representations disclose a part of the objective world *as seen from the subject's point of view*'; and, since the subject 'is aware of its own perspective on things, then by the same token, it is aware of itself as having a determinate location in the same space and time as the things it perceives' (Valaris 2008, 6). The problem with this approach is that it is not clear why I should not be able to imagine *others* occupying the same spatio-temporal location or perspective that I occupy. Since I *cannot* imagine others sharing my empirical I, the empirical I cannot be simply identified with a spatio-temporal perspective. Rather, there must be an original Sartrean awareness of myself in inner sense as radically singular and irreplaceable, underlying my awareness of the perspective on the world *that I occupy*. This radically singular self-awareness is precisely the kind of self-awareness that original consciousness of my existence in time provides: for to be aware of myself as existing in time is not *ipso facto* to be aware of the shareable, particular position in the objective time-order that I occupy. The empirical I is originally aware of itself as having no position, but as constantly fleeing any position, as time itself does. My being in time *escapes* all objectification.⁴⁵

The Sartrean approach to the transparency of the empirical I is suggested by some of Kant's remarks about awareness of myself as *existing*. Kant says that the category of existence applies to outer objects, presumably because we can subsume them under this concept: we can think of them as particular things that exist. This requires a distinction between the existences (the objects) and the existence of these existences (the category). Outer objects are particulars that have being, implying a distinction between particular beings and their being.⁴⁶ By contrast, when I am conscious of *myself* existing, I do not employ the category of existence ('here existence is not yet a category' B423n.). Presumably this is because I am not given to myself as a particular thing that can be subsumed under this concept. I am aware of existing, but not of being a particular thing that exists. All that I am, *qua* existence, is a realm of possibilities of what I can become; I am nothing determinate, not even a self-determining being, but am sheer *determinability*: 'I cannot determine my existence as that of a self-active being, rather [...] my existence always remains only sensibly determinable, i.e., determinable as the existence of an appearance' (B158n.). We should thus say not that I am a thinking thing that exists, but that '**I exist** thinking' (B420).⁴⁷ My existence

expresses an indeterminate empirical intuition, i.e., a perception [...] but it precedes the experience that is to determine the object of perception through the category in regard to time; and here existence is not yet a category, which is not related to an indeterminately given object, but rather to an object of which one has a concept, and about which one wants to know whether or not it is posited outside this concept (B423n.).

I am conscious of my existence in indeterminate, temporal intuition in a way that precedes the application of the category of existence to outer objects. My existence is not that of an intentional object of thought, but it is the possibility of standing-out (*ex-istere*), or of the emergence in time of thinking itself.⁴⁸ My existence is thus the openness (determinability) of being in time. One wants to know whether there is a thing – an object in the spatial order – that *has* this being (an object ‘posited outside this concept’), but one cannot know this at all, since one cannot know that the object given to me in outer sense is non-accidentally I (i.e. is non-accidentally identical with the existing act of thinking).

Thus, the ‘I’ of empirical apperception is not originally a referring expression used in self-ascriptions, because it does not refer to a given particular being, or even a particular perspective, in the world, but, like Sartre’s non-thetic I, is radically singular by distinguishing itself from all particulars. The transition from the universal I of <I think> to the radically singular I of indeterminate intuition in <I exist> is not the transition to a particular being encountered at a location in a spatio-temporal framework.⁴⁹ However, it is the transition to an I that *can* be situated as a particular in the world, because I can, through syntheses of the imagination under concepts of the understanding, *determine* my existence in relation to other objects.⁵⁰ As a determinate perspective in the objective spatio-temporal order, I am an object of *thought* (or cognition), not just an object of indeterminate empirical intuition. But since this self-determination is downstream of an originally *indeterminate* intuition of myself, the I that is determined is originally nothing – not even a particular position or perspective alongside others.⁵¹ Like Aristotle says of the passive intellect, the empirical I is nothing at all before it thinks (Aristotle DA, 429a23).

In thinking, the empirical I becomes something, because its existence acquires determination (KrV B157-8n.). I become aware of my thoughts’ position in a temporal order by (myself) ordering them in time. The empirical I thus is not already, as an object of (indeterminate) empirical intuition, confined within an objective, spatio-temporal framework (it is not already at a temporal location, already something), as on an Evansian account of the self-conscious subject. Rather, the transcendental I determines its temporal position itself through syntheses of the imagination (syntheses of ordering in time), in accordance with the transcendental principles of objective thought in general. There are not two acts here, but one: the empirical I’s being-determined is the same activity as the transcendental I’s act of determination. In the same way, being kicked and kicking are not two acts, but a single one, seen either from the side of agency or from the side of the patient. But in the case of kicking, the agent who kicks is distinct from the patient who is being kicked (or they are, *per accidens*, the same). By contrast, the patient or ‘empirical I’ is nothing at all before it is determined. Hence, the identity of determining and being determined reveals the (non-accidental) identity of transcendental and empirical I’s.

4. Conclusion

On the reading I have developed, we are able to say that the transcendental I is the same as the empirical I, since it is the same act of thinking of which I am conscious non-temporally in transcendental apperception, and temporally in empirical apperception. This sameness can be expressed, not by saying that the empirical I can be subsumed as an already given particular under the universal, transcendental I, but by saying that the empirical I is the self-particularizing and self-concretizing universal I. I am originally not a being, but am thinking or being (truth) itself, insofar as it is capable of emerging in time. Transparency accounts of pure apperception, by contrast, conceive of the transcendental self as originally escaping itself, leaving itself behind, in its engagement with the world. They can only understand the empirical self to be one of the objects it encounters, and thus to be distinct from it.

Although it would exceed the limits of this paper to discuss this in more detail, I think this reading of the relation between transcendental and empirical apperception puts us in a better position to understand our community with other empirical subjects. For now we can say that consciousness in general is common to, or the same across, the first-person singular, the second person, and the third person (in the first sense of 'sameness' articulated above). Reason sees itself in the other, and at the level of pure apperception, when it grasps being in the sense of being-true, sees *through* the other. But it is also, in its descent into a spatio-temporal world, conscious of itself as dividing into a community of empirical subjects and as the 'same as' each of them (in the second sense of sameness distinguished above). Thus we can say not only that there is a plurality of first persons, but that there is a first-person plurality – not only a collection of *I*'s, but a *we*: for the division of consciousness in general into individual, singular consciousnesses is internal to the transcendental I. However, I cannot say more about this here: this is a difficult topic that will have to await further study.

Notes

1. See, for instance, (Evans 1982, Allison 2004, 290f.,; Strawson 1966, 248f.. Boyle 2009) distinguishes 'an active and a passive kind of self-knowledge', suggesting that passive self-knowledge or inner sense is a kind of inner experience (Boyle 2009, 160).
2. I am grateful to Adrian Haddock for helping me formulate my claim in these terms.
3. Following the recommendation of an anonymous reviewer, I use brackets to mention thought-contents.

4. On this approach, I-thoughts have a peculiar, first-personal Fregean 'sense' or mode of presentation [*Art des Gegebenseins*] of the object they refer to. Longuenesse follows this approach when she writes, "I' and 'A' are different modes of presentation of the entity that is also referred to by the proper name, 'E.A.'. 'A' is sufficiently specified by the description: 'name that refers to whoever is currently saying or thinking 'A is F'; but 'I' needs the further specification – 'word that refers to whoever is currently saying or thinking "I am F" and whose use depends on non-thetic consciousness (of) whoever is saying or thinking "I am F"' (Longuenesse 2017, 65).
5. This is also implied by Strawson's claim that 'the concept of a person is logically prior to that of an individual consciousness' (Strawson 1966, 103). The concept of a person is that of a being in the world, and that of individual consciousness is that of the consciousness possessed by this being in the world. The first-person standpoint is thus parasitic on there being an objective world, and a person within it.
6. The identification of myself with a particular individual in the world is a self-locating thought that I can think only from within a first-person perspective. Thus, in Kantian terminology, Evans's point is that we can become aware of ourselves as elements in an objective spatial order from within inner sense – indeed, Evans is suggesting that there is bodily self-awareness in inner sense. Evans criticizes Kant for thinking that awareness of oneself as in space would require knowing ourselves as others know us – as objects of outer sense: 'The idea that I can identify myself with a person objectively construed is often mis-expressed, e.g. in terms of the idea that I realize that I am an object to others (also an object of outer sense, as Kant says: Critique of Pure Reason, B145). This misleadingly imports an ideal verificationist construal of the point' (Evans 1982, 210).
7. To identify myself with a being in the world is to situate myself within the external, objective context: 'to know what it is for $[\delta t = I]$ to be true, for arbitrary δt , is to know what is involved in locating oneself in a spatio-temporal map of the world' (Evans 1982, 211). It is unclear exactly how Evans thinks he can get self-reference just out of self-locating thought (see Sebastian Rödl's criticism of this move in his 2017, 280ff.). Perhaps Evans thinks that insofar as I am conscious of myself as occupying a particular perspective from within the world, and thus as having limits, I am conscious of myself as a particular being, because the limits are here understood as given to consciousness – or as simply encountered. But they are not encountered in the same way that I encounter outer objects. They are encountered as *my* limits – as the limits through which I engage with the world (the world would be nothing to me without them). Whereas Evans thinks of the transparent or 'purely formal' I as the limit of the entire world, and thus as lacking a perspective from within it, he thinks of the empirical *object-I* as a perspective that defines the limit of my ego-centric world. So he thinks that I encounter the empirical I as an object of reference, because the empirical I is the perspective that I encounter as a special kind of inner object. At the end of this paper, I will use Kant to criticize the Evansian view that the empirical I is a perspective, and will argue along Sartrean lines that it is nothing at all (not even a perspective).
8. See, for instance, (Byrne 2011, 2018).

9. This is also Strawson's strategy: 'If we try to abstract this use, to shake off the connexion with ordinary criteria of personal identity, to arrive at a kind of subject-reference which is wholly and adequately based on nothing but inner experience, what we really do is simply to deprive our use of "I" any referential force whatever. It will simply express, as Kant would say, "consciousness in general"' (Strawson 1966, 166).
10. This is taken from Sartre's criticism of Husserl. See fn. 19.
11. McDowell notes in his Appendix to Evans's chapter on self-identification that Evans himself became increasingly aware of the dependence of the 'fundamental level of thought' on the subjective viewpoint of the egocentrically located empirical subject. This can make it sound like he began to concede the subjective idealistic consequences of his view: 'Section 6.3 [...] gives the impression that the objective or impersonal mode of thought about space can be understood as a mode of spatial thinking organized around a framework of known objects and places – the "frame of reference". But such a mode of thinking will not be capable of achieving a higher degree of impersonality than that achieved by the subject's thought about the objects and places which constitute the frame; [...] it seems plausible that a subject's right to be counted as thinking about these familiar objects and places turns partly on his conception of the role they have played in his past life – being visited by him, seen by him, etc. [...] In that case, the seemingly objective mode of thinking about space is, after all, contaminated by egocentricity' (Evans 1982, 265).
12. Philosophers such as Richard Moran or Christine Korsgaard, who think of the first-person standpoint as the standpoint of a particular being in the world, and thus as already confined to a region within it, would not be Kantian in this sense (Moran 2001, 63; cf. Korsgaard 2009, 125f.).
13. 'Consciousness is aware of itself insofar as it is consciousness of a transcendent object' (Sartre 1960, 40). 'Every positional consciousness of an object is a non-positional consciousness of itself' (Sartre BN 1956, 40).
14. In *The Transcendence of the Ego*, Sartre puts this by saying that consciousness is 'purely and simply the consciousness of being consciousness of that object. This is the law of its existence. We should add that this consciousness of consciousness [...] is not positional, which is to say that consciousness is not for itself its own object. [...] Now we ask: is there room for an I in such a consciousness? The reply is clear: evidently not' (Sartre 1960, 40–41).
15. Sartre does say that non-thetic self-consciousness is immanent to itself, but it is immanent to itself as an act of transcendence, of going outside (e.g. Sartre 1956, 77).
16. It is a sign of the poverty of language that we have no first-person universal pronoun, and that we always speak of first-person consciousness as a first-person 'standpoint' or 'perspective'. Both of these terms suggest that we are conscious from some place (standpoint) or position (perspective) within the world.
17. This point is helpfully made by Jean-Philippe Narboux in a recent essay on self-consciousness (Narboux, unpublished).
18. Sartre thus would not agree with Moran's claim that from 'within the first-person perspective' (non-thetic consciousness), I acknowledge that the 'fact believed and the fact of one's belief are two different matters' (Moran 2001, 62).

19. Although it is not my intention to interpret Sartre here, I do wish to distinguish the view I am calling Sartrean from other views that bill themselves as Sartrean. Matthew Boyle has argued that non-thetic self-consciousness is transparent to the world because it is entirely directed outwards, at the objects of consciousness. But he thinks that although the mental act of being intentionally directed at the world initially is not an object of consciousness, I am implicitly aware of it as the 'manner' in which I apprehend the object: '[the subject] shows an awareness, implicit but open to reflective articulation, of the specific kind of relation in which she stands to the object of her representation' (Boyle, unpublished, 27). Richard Moran similarly argues that there is an unthematized awareness of being committed to the truth of what I think when I am non-thetically aware of my judgments (Moran 2001, 84). The transition from the outward-looking thought $\langle p \rangle$ to the inner fact, $\langle I$ believe $p \rangle$, according to Moran, would not be legitimate if I did not already implicitly see myself (as an empirical subject) playing a role in the determination of what I believe through the exercise of my rational agency (Moran 2012, 3). Thus, on both Boyle's and Moran's views, the I of non-thetic self-awareness – understood relationally, as my attitude taken towards an object – can become an object of thetic self-awareness (see also Longuenesse 2017, 47). This means that the self of non-thetic self-awareness is not fully transparent in Sartre's sense, for the subject is (even if only implicitly) aware of the manner in which she looks outwards, towards the world. Since she implicitly sees the mode of apprehension, in addition to that which she sees through it, her mind cannot be said to be aware only of what is outside, but is also aware of something inner (a 'manner of apprehending' the object, or an exercise of rational agency). And it is this implicit awareness of something inner – not merely the awareness of something outer – that, on these views, licenses the explicit self-ascription of mental states in *thetic* self-consciousness. Boyle's and Moran's views are empirical variants of the Husserlian view that Sartre rejects, according to which the transcendental I is 'so to speak, behind each consciousness, a necessary structure of consciousness whose rays (*Ichstrahlen*) would light upon each phenomenon presenting itself in the field of attention' (Sartre TE, 37). As Sartre argues, this implicit awareness of the structure through which I am aware of the world would make consciousness 'personal': it would introduce a given character of the subject of awareness into her consciousness of the objective world. (In Husserl, this character is innately given, while in Boyle and Moran it is empirically given.) But this would mean that it destroys the objective character of what she is conscious of: the world would become the world *as it appears to her*, through her manner of apprehending it. In Sartre's words, Husserl's transcendental ego would 'divide consciousness; it would slide into every consciousness like an opaque blade', and would thus be 'the death of consciousness' (Sartre TE, 40).
20. It is important to distinguish these two senses of 'identity', because otherwise these two statements would be contradictory: 'I am what I am not' and 'I am not what I am'. It does not help to insist that I am a contradiction, for which I is it that asserts itself to be a contradiction? We would have to posit a *third* I that both is identical with the posited ego and is not identical with it.

21. In this paper I am primarily concerned with the sameness of the transcendental and empirical I's of theoretical cognition. As Kant explains in the introduction to the third *Critique*, theoretical cognition (and therefore theoretical philosophy) comprises not just perceptions, but also actions (KU 20:200–1, 200n.). So the empirical I of theoretical cognition is both an I that perceives and an I that acts. In practical cognition, peculiar difficulties arise with regard to the relation between the noumenal self (I of moral freedom) and the phenomenal self (the I that acts in nature), which I will set aside for my purposes here.
22. This section is heavily indebted to Stephen Engstrom's reading of Kant's Copernican turn in his article 'Knowledge and its Object' (Engstrom 2017).
23. For a different reading of this passage, according to which false judgments relate to an object *insofar* as they are false, see (Tolley 2011, 204).
24. See (Allison 2004, 87–88; Longuenesse 1998, 82).
25. The core case of judgment is assertion, which presents a thought as 'actual (true)' (A74/ B100). On problematic judgment, see fn. 25.
26. Kant obviously acknowledges problematic judgments that do not involve an awareness of the (sufficient) grounds of the truth of the judgment – and thus are not assertoric. However, *problematic* judgments are not the same as *merely problematic* judgments, which strictly speaking are not judgments at all, but mere thoughts. Problematic judgments, such as opinions, still count as judgments because they have 'connection with truth which, although it is not complete, is nevertheless more than an arbitrary invention' (A822/B851). Problematic judgments can 'grow up' to become assertoric and apodeictic, i.e., can be 'gradually incorporated into the understanding' (A76/B101). By contrast, *merely* problematic judgments do not make a truth claim, but instead think of a thought as merely logically possible (i.e., involving no contradiction), and thus cannot become assertoric.
27. I am grateful to an anonymous reviewer for helping me clarify this.
28. Jennifer Hornsby has called this an 'identity theory of truth' (Hornsby 1997).
29. 'Synthesis alone is that which properly collects the elements for cognitions and unifies them into a certain content' (A77-8/B103). This can be understood as indicating that content is itself a holding together or synthesis of elements of cognition. This would require further elucidation, but it exceeds the limits of this paper to provide it here. See (Tolley 2011) for a Fregean reading of Kant that distinguishes content and act (force).
30. Notice that being in the sense of 'reality' or 'existence' (categorical being) is not the same as thinking, since I can think what is not real, or what does not exist. It is only being in the sense of truth (objective validity), which is 'higher' than categorical being, that is the same as thinking (judging).
31. To echo Aristotle: 'When thought has become each thing ... its condition is still one of potentiality, ... and thought is then able to think itself' (Aristotle DA, 429b6-9).
32. 'Thinking, taken in itself, is merely the logical function and hence the sheer spontaneity of combining the manifold of a merely possible intuition' (B428).
33. This metaphor of 'filling' empty forms with content can be misleading, since as discursive forms the logical functions of judging are not to be understood on the model of empty forms of intuition that get filled with matter. Kant distinguishes the way the transcendental I (or capacity for judgment, as an analytic-universal) contains all things from the way time contains all that

happens (as a form of intuition), and from the way God contains all things (as a synthetic-universal) (KU §76). Crucially, there is a sense in which the function of judging itself is still empty, even when it has been filled with 'transcendental content', since it remains a capacity that can be employed in infinitely many other judgments. This emptiness of logical functions of judging (qua capacities) is a way of characterizing the finitude of the discursive intellect.

34. 'Objective validity and necessary universal validity are reciprocal concepts' (P 4:298; see also A104-5).
35. I am grateful to an anonymous reviewer for helping me clarify this point. Notice that on this reading, thought is not initially grasped as the thinking of a being situated within an objective context. Contrary to the Evansian approach to transparency, thinking is not an occurrence in a subject located within an objective, spatio-temporal world, rather the spatio-temporal world, including the time-determinations of the schematism chapter, are internal to thought. Truth therefore will not be understood as truth-at-a-context, rather all contexts will be internal to truth. And most importantly for our topic, self-consciousness will not be consciousness indexed to an individual in the world, rather the singularity of the thinker will be internal to 'consciousness in general', the I of pure apperception.
36. I am borrowing from Sebastian Rödl's language in his new book, *Self-Consciousness and Objectivity* (Rödl 2018, 1ff.).
37. In other words, as I argued in the last section, the transcendental I (reason) thinks only itself. An anonymous reviewer has expressed skepticism about my construal of reason or apperception here as 'fundamentally generic or intersubjective rather than individual'. But if the primary truth-bearers are acts of judging (not act-independent contents), and if truths can be shared with others in communication, then surely self-conscious acts can be shared. This means that the I is shared, since the transcendental I is just the unity of such acts of judging. We should not dogmatically assume that reason and its exercises are powers or attributes attached to individual subjects. Instead, if the argument for the priority of transcendental over empirical apperception is valid, we should think of individual subjects as first made possible by the (universally shared) capacity for knowledge (i.e., by reason, or the transcendental I).
38. A species is only partially homogeneous with a genus. For instance, an ostrich is an animal, but it is *more* besides that. By contrast, all species are fully the same as the <I think>, since there isn't anything more (anything additional, not contained in the <I think>) that they can be. All species determinations of genera are therefore entirely internal to <I think>. So although the <I think> is not a particular concept (or genus), but the form of all concepts in general, one can nevertheless think of particular concepts as specifications of it.
39. According to Allison, inner sense does not 'relat[e] representations to objects' but instead 'makes these representations themselves into (subjective) objects, which it cognizes as the contents of mental states' (Allison 2004, 278–9). This would make my inner states into objects given to me; but Kant appears to be saying that representations cannot be made into objects. We can only become conscious of them through the acts of representing outer objects.
40. Valaris emphasizes this aspect of inner sense in his interpretation, to which I am heavily indebted (Valaris 2008).

41. Contrary to Boyle, who restricts the 'objects of inner sense' to 'sensations, appetites, and other kinds of mental "affection"' (Boyle 2009, 160), I think it is clear that Kant thinks we can become empirically aware not just of our passive states, but also of our acts of thinking and judging: 'I can say that I as intelligence and thinking subject cognize my self as an object that is thought, insofar as I am also given to myself in intuition' (B155).
42. In inner sense, Kant says that 'I merely represent the spontaneity of my thought, i.e., of the determining, and my existence always remains only sensibly determinable, i.e., determinable as the existence of an appearance' (B158n.).
43. In *Self-Consciousness*, Sebastian Rödl argues that all knowledge is *either* 'receptive knowledge' or 'spontaneous knowledge', and that first-personal self-knowledge is spontaneous knowledge. But if my reading of self-affection is accurate, it is both receptive *and* first-personal or spontaneous (see, for instance Rödl 2007, 144f.).
44. Strictly speaking the transcendental I appears to be 'nothing' in a sense that precedes the 'concept of an object in general', which Kant says underlies the distinction between 'something' and 'nothing', and thus precedes both a table of something and the 'table of nothing' (A290/B346). As a logical sort of 'nothing', the transcendental I is the concept of the 'original-synthetic unity of apperception' that precedes the 'objective unity of apperception' and thus precedes the concept of an object (B131ff.). The 'nothingness' of the empirical I, like that of the logical I, also does not belong on the 'table of nothing' at A292/B348, because the empirical I, as an object of indeterminate empirical intuition, is not an object of determinate empirical intuition, and so does not fall under the 'concept of an object in general'. I am thankful to Addison Ellis for pressing me to think about these different senses of 'nothing'.
45. As I will clarify in the following, this is not to deny Valaris's point that I can become aware of my particular perspective in space and time, and can thus situate myself in the world as an appearance alongside other appearances, through syntheses of the imagination in inner sense. I only mean to deny that this can be an exhaustive account of inner sense.
46. There is also a sense in which the object is the same as its being, since objects are homogeneous with the concepts they can be subsumed under (as argued above). But the sameness of the object and its concept in that sense is merely formal. Since this formal object is the same across different material objects (identity in difference), there is still a distinction between what is the same (formal object: being) and what is different (material objects: beings).
47. This contrasts with Longuenesse's reading, according to which I refers to a thing that exists even at the level of transcendental apperception: 'We just learn to use "I" to refer to ourselves insofar as, necessarily, in thinking we ascribe thinking to ourselves, the individual currently engaged in the act of thinking, and aware of thinking by perceiving the fact that we think' (Longuenesse 2017, 89). See also Kitcher, who interprets empirical apperception as a kind of self-ascription of mental states (Kitcher 2011, 124).
48. This is an allusion to Heidegger's analysis of the etymology of existence ('ek-sistence') as a standing-out (Heidegger 1998, 147f.).
49. I am grateful to Sebastian Rödl for pointing out to me the distinction between 'particular' and 'singular' in this context.

50. Valaris helpfully shows how even my spatial perspective is something that I become aware of in inner sense, since the juxtaposition of things in space is an act of imaginative synthesis (placing one thing alongside another) and hence available to empirical self-awareness (Valaris 2008, 8f.).
51. Kant indicates that these acts of self-determination presuppose an indeterminate intuition of my existence in the following passage: 'The **I think** expresses the act of determining my existence. The existence is thereby already given, but the way in which I am to determine it, i.e., the manifold that I am to posit in myself as belonging to it, is not thereby given' (B157n.).

Acknowledgments

I have presented previous versions of this paper at three events. The first was a conference on *Introspection, Self-Consciousness, and Self-Knowledge* at the University of Nebraska, Lincoln, in May 2015. The second was a conference on *Transparency and Self-Consciousness* at Ryerson University in Toronto in May 2018. And the third was a colloquium talk at the University of Leipzig, Germany in June 2018. Thanks to all the audiences at these events for their challenging questions and astute comments. Special thanks to Irad Kimhi, Sebastian Rödl, Matthew Boyle, Adrian Haddock, David Sussman, and Stephen Engstrom for helpful conversations directly related to the topic of this paper. I also thank two anonymous reviewers for their extensive and detailed comments, many of which led me to rewrite the latter half of this paper.

Notes on contributor

Alexandra M. Newton received her PhD in philosophy at the University of Pittsburgh, where she wrote a dissertation on Kant on logical form. She has worked as a 'wissenschaftliche Assistentin' at the University of Leipzig in Germany and is currently employed as an assistant professor of philosophy at the University of Illinois at Urbana-Champaign. Her interests include all topics discussed by Kant's works, with an emphasis on the role of self-consciousness in logic, metaphysics, practical philosophy, and aesthetics

ORCID

Alexandra M. Newton  <http://orcid.org/0000-0002-7491-9728>

References

- Allison, H. 2004. *Kant's Transcendental Idealism*. New Haven: Yale University Press.
- Aristotle. 1984. "De Anima" (DA). Translated by J. A. Smith. In *The Complete Works of Aristotle*. 1 Vol. edited by J. Barnes. Princeton: Princeton University Press.
- Boyle, M. 2009. "Two Kinds of Self-Knowledge." *Philosophy and Phenomenological Research* 78 (1): 133–164. doi:10.1111/phpr.2008.78.issue-1.
- Boyle, M. unpublished. "Transparency and Reflection."

- Byrne, A. 2011. "Transparency, Belief, and Intention." *Proceedings of the Aristotelian Society Supplementary Volume* 85 (1): 201–221. doi:10.1111/j.1467-8349.2011.00203.x.
- Byrne, A. 2018. *Transparency and Knowledge*. Oxford: Oxford University Press.
- Engstrom, S. 2013. "Unity of Apperception." *Studi Kantiani*, Fabrizio Serra Editore, 26: 37–54.
- Engstrom, S. 2017. "Knowledge and Its Object." In *Kant's Critique of Pure Reason: A Critical Guide*, edited by J. O'Shea, 28–45. Cambridge: Cambridge University Press.
- Evans, G. 1982. *Varieties of Reference*. Edited by J. McDowell. Oxford: Oxford University Press.
- Heidegger, M. 1998. "On the Essence of Truth." In *Pathmarks*, edited by W. McNeill, 136–154. Cambridge: Cambridge University Press.
- Hornsby, J. 1997. "Truth: The Identity Theory." In: *Proceedings of the Aristotelian Society* 97: 1–24, Aristotelian Society was held at Birkbeck College, London.
- Kant, I. 1992. Jäsche Logic (JL). In: *Lectures on Logic*. Translated and edited by J. M. Young. Cambridge: Cambridge University Press.
- Kant, I. 1998. *Critique of Pure Reason* (KrV). Translated and edited by Guyer, Paul and Wood, Allen W. Cambridge: Cambridge University Press.
- Kitcher, P. 2011. *Kant's Thinker*. Oxford: Oxford University Press.
- Korsgaard, C. M. 2009. *Self-Constitution. Agency, Identity, Integrity*. Oxford: Oxford University Press.
- Longuenesse, B. 1998. *Kant and the Capacity to Judge*. Princeton: Princeton University Press.
- Longuenesse, B. 2017. *I, Me, Mine. Back to Kant, and Back Again*. Oxford: Oxford University Press.
- McDowell, J. 1994. *Mind and World*. Cambridge, MA: Harvard University Press.
- Moran, R. 2001. *Authority and Estrangement. An Essay on Self-Knowledge*. Princeton: Princeton University Press.
- Moran, R. 2012. "Self-Knowledge, 'Transparency', and the Forms of Activity." In *Introspection and Consciousness*, edited by D. Smithies and D. Stoljar, 211–238. Oxford: Oxford University Press.
- Naroux, J.-P. unpublished. "Is Self-Consciousness Consciousness of One's Self?"
- Rödl, S. 2007. *Self-Consciousness*. Cambridge, MA: Harvard University Press.
- Rödl, S. 2017. "The First Person and Self-Knowledge in Analytic Philosophy." In *Self-Knowledge. A History*, edited by U. Renz. Oxford: Oxford University Press: 280–294.
- Rödl, S. 2018. *Self-Consciousness and Objectivity. An Introduction to Absolute Idealism*. Cambridge, MA: Harvard University Press.
- Sartre, J.-P. 1956. *Being and Nothingness (BN)*. Translated by H. E. Barnes. Gallimard. New York: Washington Square Press.
- Sartre, J.-P. 1960. *The Transcendence of the Ego. An Existentialist Theory of Consciousness (TE)*. Translated by F. Williams and R. Kirkpatrick. New York: Hill and Wang.
- Sartre, J.-P. 1970. "Intentionality: A Fundamental Idea of Husserl's Phenomenology." *Journal of the British Society for Phenomenology* 1 (2): 4–5. doi:10.1080/00071773.1970.11006118.
- Strawson, P. 1966. *The Bounds of Sense. An Essay on Kant's Critique of Pure Reason*. London: Routledge.
- Tolley, C. 2011. "Kant on the Content of Cognition." *European Journal of Philosophy* 22 (2): 200–228. doi:10.1111/j.1468-0378.2011.00483.x.
- Valaris, M. 2008. "Inner Sense, Self-Affection, and Temporal Consciousness in Kant's Critique of Pure Reason." *Philosophers' Imprint* 8 (4): 1–18.
- Wittgenstein, L. 1958. *Blue and Brown Books*. New York: Blackwell Publishers.
- Wittgenstein, L. 2001. *Tractatus Logico-Philosophicus (Tractatus)*. New York: Routledge.