

*The Achilles after the Paralogisms*¹

Abstract: According to a broad consensus, the Paralogisms section of the *Critique* is understood to reject substantial conceptions of the subject, and specifically so-called “Achilles” arguments for these conceptions proposed by such rationalist philosophers as Wolff, Baumgarten and Mendelssohn. But I argue the consensus is wrong, and Kant’s target in the Paralogisms is limited to empirical versions of the Achilles. I argue Kant himself accepts a substantial view of the self, as well as a version of the Achilles.

In the B-edition of the first *Critique*, Kant appends a section (“Refutation of Mendelssohn’s Proof for the Permanence of the Soul” (B413-418)) specifically intended to dismantle an argument proposed in the Second Dialogue of Moses Mendelssohn’s *Phaedon* (2007), which seeks to establish the simplicity and incorruptibility of the soul. Kant famously characterizes this type of argument as (A351):

the Achilles of all dialectical inferences of the pure doctrine of the soul. It is no mere sophistical play, contrived by a dogmatist in order to impart to his assertions a superficial plausibility, but an inference which appears to withstand even the keenest scrutiny and the most scrupulously exact investigation.

Kant takes Mendelssohn’s argument to exemplify a foundational theme in the rationalist tradition of establishing the existence of a simple and unified human soul. In particular, the Achilles has since come to denote any argument that proceeds from premises concerning the presence of a certain “unity” in mental goings-on—e.g. thoughts, inferences or visual or qualitative experiences—to the conclusion that there must exist a self that is in an important sense simple, unified or otherwise part-less.² Thus in J.P. Schachter’s description, an Achilles argument (Schachter 2008, p. 177; italics original):

[h]ypostasizes [...] a ‘simple’ or indivisible substance which, by virtue of being simple, is then further inferred to be immaterial. [...] it attempts to do this on the basis of experiential data consisting of various kinds of ‘unity’ discoverable in our experience, namely (a) the unity of multiple distinct perceptions in a single subject, (b) the unity of multiple ideas in a judgment, and (c) the unity present in and underlying self-awareness

¹ For clarity of presentation, works by Kant, Leibniz, Wolff and others are cited by abbreviation, an overview of which is appended to the bottom of this document.

² Achilles arguments are not prominent merely in early modern and rationalist literature, but have seen an upsurge in interest in contemporary philosophy of mind as well. See e.g. Rovane 1998, 2012 and Burge 1996 for the claim that inference requires an absolutely simple subject, and Barnett 2008 for the argument that an intuition of the Achilles form underlies several important contemporary thought experiments in the philosophy of mind (Searle’s Chinese Gym, e.g.). See also Barnett 2010 for a positive argument that consciousness must be simple.

[...] For an argument to satisfy the Achilles template, it is not sufficient that it somehow make reference to the unity of consciousness; rather, it must use the unity of consciousness as the ground for an inference to a simple substance that is the subject of experience.

Kant famously rejects the Achilles argument as a “paralogism”: an argument that is bound to compel us due to the structure of reason, but the appeal of which is ultimately illusory and which we can learn to resist (A341/B399).³ As some commentators have noted, it can seem surprising that Kant rejects the Achilles. This is for two reasons.

First, a unified conscious subject plays an important role in Kant’s critical account of cognition. Consider, for example, Kant’s powerful suggestion of the importance of unity in subjectivity (A354; italics mine):

It is obvious that if one wants to represent a thinking being, one must put oneself in its place, and thus substitute one’s own subject for the object one wants to consider (which is not the case in any other species of investigation); and it is obvious that we demand *absolute unity of the subject of thought*.

After the initial suggestion in this passage of a relation between representing thought and one’s own subjectivity, Kant further seems to positively claim that an absolute unity of the subject is required for understanding another as a thinker. Just so, Stephen Engstrom’s interpretation of the Original Synthetic Unity of Apperception grounds “the generality of representation that all concepts share in common” in “the identity, or oneness-in-many, of the primitive act of consciousness itself” (Engstrom 2013, p. 7). Of this “primitive act of consciousness” Engstrom further writes (Engstrom 2013, fn. 13.):

This act is also simple: Kant describes the subject’s “discursive consciousness” of this act, or the “pure apperception of its mental action” as “simple”, adding, “The I of reflection contains no manifold in itself and is in all judgments always one and the same” (Anth 141; cf. B135, B138).⁴

Second, Kant’s pre-critical metaphysical thinking—as attested by metaphysical lectures, *Reflexionen* and notes dated to the middle and late 1770s, i.e. close to the publication of the first

³ For discussion of Kant’s notion of a paralogism, see Harper 2008, pp. 235-237, pp. 244-245. For Kant’s belief that paralogisms stem from the nature of reason itself, Michelle Grier 2001. See also Grier 1993.

⁴ Compare also Sassen 2008 (p. 216, fn. 6): “Kant himself advanced at least a portion of an argument like the one we find in the Second Dialogue of the *Phaedon* in the Transcendental Deduction of the Categories in the second or B-edition of the Critique (B130-40).”

edition of the *Critique* in 1781—seems to positively endorse ontologically substantial accounts of the “I” that Achilles arguments are typically intended to establish. So Kant writes in a note dated to the period 1775-1777: “The understanding itself (a being that has understanding) is simple. It is substance” (R 4785, 17:707).⁵ This apparent contrast with Kant’s critical view in the Paralogisms has led some commentators to suggest that Kant’s thinking undergoes a *volte-face* around 1780, accepting something like the Achilles until into the 1770s before rejecting it in the *Critique*. As Pierre Keller writes (Keller 1998, Chapter 8, n. 1):

Kant did not free himself from commitment to the project of rational psychology until quite late in his career. Kant still thought of the subject of thought as theoretically knowable in the substantial terms suggested by the four basic paralogisms that he identifies in rational psychology into the middle of the 1770s. His discovery of the fallacies involved in inferences from self-consciousness to substantive claims about the nature of the self or the soul was the last important innovation in his thinking prior to the publication of the *Critique* in 1781.⁶

But against the consensus, in this paper I will argue for the following claims:

(1) The Continuity of the Paralogisms: Appearances in the Paralogisms section notwithstanding, Kant’s critical philosophy is consistent with his late pre-critical work in holding to a conception of the self as a simple, noumenal substance.

(2) The Specificity of the Paralogisms: Kant’s criticisms in the Paralogisms of the Achilles are targeted at specific, identifiable versions of such arguments advanced by Kant’s rationalist predecessors, as opposed to arguments of the Achilles variety *simpliciter*.

Taken together, I suggest **(1)** and **(2)** provide room for a third claim that can seem surprising:

(3) Kant’s Critical Achilles: Kant’s critical thought offers a version of the Achilles argument.

If **(1)-(3)** are true, it reorients our understanding of Kant’s relation to the German rationalist tradition criticized in the Paralogisms. On my suggested view, it is Kant’s advocacy of a change in the nature of the Achilles argument as advanced by previous German rationalists that

⁵ Cited at Messina 2014, p. 36.

⁶ Similarly, Corey Dyck considers Kant’s metaphysical writings from the 1770s one of the main targets of Kant’s critical thinking (Dyck 2014, pp. 10-11). By contrast, recent work by Julian Wuerth (e.g. 2014) represents the opposite tendency, arguing for continuity between Kant’s metaphysical and critical thought.

has confused commentators into thinking of Kant as altogether dismissive of such arguments.⁷ In particular (in terms that will become clear), Kant dismisses a *broad* and (for Kant, associatedly) *empirical* Achilles in favor of a *narrow* and *pure* Achilles.⁸

I will proceed as follows. In §1 I introduce Lennon and Stainton 2008's helpful distinction between "narrow" and "broad" Achilles arguments. In §2 I turn to Kant's critique of the Achilles in the Paralogisms section of the *Critique*. In §3 I raise a dilemma for the consensus understanding of the Paralogisms, suggesting the common reading cannot be held together with a plausible reading of Kant's metaphysical writings. In §4 I resolve this dilemma by suggesting a reading of the Paralogisms as tailored to address specific arguments proposed by Kant's predecessors Alexander Baumgarten and Christian Wolff. But my argument is not merely textual. In §5 I associate my reading of the Paralogisms with a positive philosophical understanding of Kant's view of the relation between self and self-consciousness, which centers on a version of the Achilles. In §6 I discuss objections.

Before proceeding, I want to register a note on the philosophical substance at stake in defending claims **(1)-(3)**. It may seem of limited philosophical interest to know whether Kant endorses an Achilles argument, or whether Kant accepts an account of the self as metaphysically substantial and simple. But my suggestion is that this impression is wrong and that there are

⁷ For an excellent exposition that similarly sees important vestiges of rationalist views of the self in Kant's metaphysics of mind, see Tester 2014, e.g. p. 12:

The tendency to overlook the rationalist background of Kant's discussion and the profound influence of rationalism upon his thought have led interpreters to misunderstand the actual arguments that Kant is targeting in his discussion of rational psychology and more importantly to overlook aspects of Kant's discussion that reveal a commitment to positive metaphysical doctrines very similar to those proposed by the rational psychologists and continuous with the rationalism of his pre-critical writings.

For more work sympathetic to significant linkages between Kant's critical thinking and his rationalist predecessors, see Wunderlich 2001, 2005, Proops 2010, Watkins 2005, Wuerth 2006, Wuerth 2010a, 2010b, 2014, Dyck 2008, 2011, 2014.

⁸ In this regard my reading of Kant will depart markedly from the consensus (the *locus classicus* of which is at Strawson 1966, pp. 155-170). On the standard view the Paralogisms are read as a rejection of the sort of "dogmatic metaphysics" that proceeds in independence from experience. But I argue the opposite is true, and a form of *empiricism* is among Kant's targets.

significant philosophical and explanatory benefits in appreciating substantial language in Kant's critical views.⁹ Central to Kant's thinking is the claim that the apperception expressed by "I think" lies at the heart of the mind, and must be capable of accompanying all representations. In this context, Kant frequently speaks of the "form" and "unity" of representations. But it has not been clear how to understand this language, and what might justify the universal application of apperception to mental states. Here I will argue that apperception is central to Kant's Achilles argument, and more specifically, is constitutive of a type of substance. If this is correct, I suggest it gives us a new purchase on the strength of Kant's apperception claim. If apperception is constitutive of a substance, then it is intelligible how representations, as accidents inhering in this substance, must conform to the substance's apperceptive nature. Apperception, we might say, provides a way for accidents to inhere in the substance it enforms, thus demanding conformity of the accidents to the substance's apperceptive nature.¹⁰ Moreover, this construal also adds an axis to the debate about the tenability of Kant's account of apperception, viz. the role of apperception in accounting for the unity or simplicity of the conscious mind. Whether unity and simplicity are important properties of consciousness, and whether apperception is plausibly constitutive of them, can accordingly be used to motivate arguments for and against Kant's doctrine of apperception.

1. *Achilles: Narrow and Broad*

In their illuminating analysis of the structure of Achilles arguments, Lennon and Stainton 2008 distinguish "narrow" and "broad" versions of the argument.¹¹ Lennon and Stainton's

⁹ For an excellent exposition of the explanatory benefits of substantial language in understanding Kant's view of the intellect, see Boyle (Forthcoming).

¹⁰ Of course, this is not yet to *justify* thinking of apperception as fundamental in this sense.

¹¹ Lennon and Stainton further note a "narrowest" Achilles argument characterized by one particular *reductio* to support P2, and a "broadest" Achilles argument that moves from C1 to conclusions beyond the nature of the soul (for example, contemporary philosophers of consciousness interested in feature-binding, non-modular mental unity or phenomenal integration might argue from premises about the unity of mental phenomena to a variety of conclusions, such as empirical hypotheses concerning the faculties that a

distinction turns on whether the Achilles seeks to establish the limited conclusion of the existence of a simple and unified soul (narrow) or whether this conclusion in turn serves as a premise for stronger claims about the nature of the soul, such as its immaterialism or immortality (broad).¹²

<p><i>Narrow Achilles</i> P1: Unification of representations takes place P2: Only a simple, unified substance can unify representations. C1: The human soul or mind is a simple unified substance.</p>	<p><i>Broad Achilles</i> P3: If the human soul/mind is a simple, unified substance, then it is not material P4: If the human soul/mind is a simple, unified substance, then it is immortal. C2: The human soul/mind is immaterial and immortal.</p>
---	--

It is not coincidental that Stainton and Lennon focus on immateriality and immortality as conclusions for the broad Achilles. For purposes of understanding Kant’s criticism of the Achilles below, it is relevant that the Achilles was typically invoked to establish broadly theological conclusions about the incorporeality and imperishability of the soul.¹³ It is helpful to very briefly sketch such arguments. The immateriality of the soul is frequently taken to follow from its indivisibility. On these arguments, spatiality is intrinsically tied to indefinite divisibility. In turn, material existence conforms to spatiality. It follows that if the soul is not divisible, it cannot be material. In turn, immateriality is thought to imply imperishability. Here the thought is that destruction must be a matter of dissolution. An immaterial, indivisible soul cannot be dissolved. Accordingly, the broad Achilles concludes, the soul as simple substance is indestructible and persists after death.¹⁴

In addition to the distinction between narrow and broad forms of the Achilles it is also helpful to distinguish between two archetypes of the argument. One of the earliest Achilles arguments is arguably found in the work of the neoplatonist Plotinus (EA IV 7.6; italics added):

If something is going to perceive anything, it *must itself be one* and must apprehend by the same [sense], both if several impressions are [perceived] through many sense organs,

developed cognitive science will posit.) Here I will not be concerned with arguments of either the narrowest or broadest Achilles varieties.

¹² For these two versions, see Lennon and Stainton 2008, p. 3, Lennon and Stainton 2008, p. 5.

¹³ This, for example, is the structure of the Mendelssohnian argument that is Kant’s original example of an Achilles.

¹⁴ For trenchant assertions of both the infinite divisibility of matter and the imperishability of substance, see Leibniz (AG 79).

or many qualities [are perceived] in one object, or if one sense organ [perceives] a complex object, for example, a face. For there isn't one [perception] of the nose and another of the eyes, but one identical [perception] of all of them together. And if one [sense-object] enters through the eyes, and another through the hearing organ, there must be some one place which they both go. Otherwise, how could we state that they are different from each other, if the sense-objects did not all come together to one and the same place? Therefore, this place must be like a center point, and the perceptions coming from all places, like the lines coming from the circumstances of the circle, must terminate here. And the thing that apprehends these perceptions must be of this sort, really one.

Close reading reveals that in this brief paragraph Plotinus manages to raise three different grounds on which to mount an Achilles argument. The first takes as its experiential datum the unity of a perceived object, specifically a face. As Plotinus suggests, a perceiver is not separately aware of the eyes, the ears and the nose, but rather perceives them as aspects of a unified face. The second of Plotinus' arguments invokes the unity of sense information deriving from different modalities, such as seeing and hearing. Plotinus' third argument expands on the argument from different modalities by drawing on the ability of subjects to compare and contrast the various representations issuing from them. As Plotinus writes, "how could we state that they are different from each other, if the sense-objects did not all come together to one and the same place?" For Plotinus, these three considerations serve the same conclusion that there must be a "thing" that is in an important sense simple.

But as helpfully discussed by Coope 2013, Henry 2008 and Cory 2014, there is a different Achilles in the vicinity, which departs not from a unity found in experience, but rather a unity found in a distinctive type of reflection. The second version of the Achilles argument is provided by Proclus (the argument is also expressed in the neoplatonist work *Book of Causes*, which is in part based on Proclus' work (Coope 2013, p. 8)). As Ursula Coope writes, for example: "Proposition 7 in the *Book of Causes* (based on proposition 15 of Proclus' [ET]) asserts that since an intelligence 'reverts upon its essence' it must be an 'undivided substance'" (Coope 2013, p. 8).

<p><i>Plotinian Achilles</i> P1: Unification of representations takes place P2: Only a simple, unified substance</p>	<p><i>Proclean Achilles</i> P1: Unity characterizes reflective consciousness P2: Only a simple, unified substance</p>
--	---

can unify representations. C1: The human soul or mind is a simple unified substance	can possess this unity C1: The human soul or mind is a simple unified substance.
--	---

For my purposes, there are two major distinctions between these two arguments. The first distinction concerns P1. The Plotinian argument departs from a process of “unification” among representations. The productive character of this premise renders salient a grounding of the existence-claim in a quasi-empirical epistemology, on which the subject introspects the product of what is subsequently inferred to be a process of unification (as is the case, e.g., in Plotinus’ observations concerning the awareness of a face).¹⁵ By contrast, the Proclean argument departs from a “unity” in reflective consciousness which carries no productive associations. In turn, this renders plausible that the claim of this unity is not grounded in introspection, but can be understood purely conceptually from the notion of reflection.¹⁶

The second distinction concerns the grounds for P2. The Plotinian argument seems to depend on an explanatory inference on which only a hypostasized simple substance can serve to explain the phenomenon recorded in P1. By contrast, the Proclean argument may be understood as giving a *constitutive* gloss on the grounds for P2. The simple substance is not *inferred* to explain the reflective unity recorded in P1. Rather, the claim is that a substance is constituted by the sort of unity implicated in reflective consciousness. This is indeed Proclus’ own intended development of the argument. In Coope’s characterization, “Proclus [...] claims that reversion to oneself is cognition of oneself ([CT], ii. 286. 32) and that in reverting on itself a thing *constitutes itself* as the kind thing it is (so that all and only those things that are revert on themselves are self-constituted, *authupostata*)” (Coope 2013, fn. 34; italics mine). To be sure, at this point it is unclear what is supposed to motivate a constitutive conception of the relation between reflection

¹⁵ To be sure, the connection is not compulsory. A Plotinian argument could, e.g., depart from *a priori* premises concerning a process of unification. But as I discuss below in regards the influential Wolffian variety of rational psychology, in fact the premise is typically understood as relying on empirical grounds.

¹⁶ If the unity of Proclean reflection is intended to be *a priori* this is not intended to rule out to the role of the cogito in arguments for the unity of reflective self-consciousness. Although I cannot here defend this as an analysis of the cogito, the suggestion is that the cogito merely constitutes thought acting out its own nature: displaying its unity to itself, rather than observing a unity in some further state of consciousness. This is in line with the structure of the ‘Merian’ cogito in §4 discussed below.

and a simple subject, or in Proclean vocabulary, the relation between reflection and “self-constitution”. I address this issue in §5 below.

2. *Kant’s Critique of the Achilles in the Paralogisms*

The consensus on Kant’s rejection of the Achilles is grounded in the Paralogism section of the *Critique*, in which Kant seems to diagnose and disarm the appeal of arguments close to the Achilles. Consider the First and Second Paralogisms as stated in the A-edition (A348, A351s; emphasis original)¹⁷

I, as a thinking being, am the **absolute subject** of all my possible judgments, and this representation of myself cannot be used as predicate of any other thing.
Therefore I, as thinking being (soul), am substance.

A thing whose action can never be regarded as the concurrence of many acting things is **simple**.
Now the soul, or the thinking **I**, is such a thing. Therefore, etc.

Cursorily stated as they are, these two arguments are clearly intended to invoke the two conjuncts of the conclusion of an Achilles: the status of the self as substantial and simple.¹⁸ In particular, the First Paralogism infers to the existence of the subject as substance on grounds of its peculiar nature as being essentially the subject of predication (I come back to this argument in detail below). The Second Paralogism comes close to the Plotinian paradigm of an Achilles argument, by arguing from an *explanandum* of an action—in this case, thinking—that brooks no multiple subjects to an *explanans* of a simple subject. Kant further illustrates this type of thought in the famous “verse argument” (MH 28:44):

¹⁷ In the A-edition, Kant provides distinct sections treating inferences to the status of the soul as a substance (First Paralogism); inferences to the simplicity or unity of this substance (Second Paralogism); inferences to the persistent identity or personality of the self (Third Paralogism); and a treatment of the ideality of the objects of experience (Fourth Paralogism). Perhaps in light of its divergent subject matter as concerned with objects rather than the self, discussion of the Fourth Paralogism in the B-edition is moved to the Refutation of Idealism. For discussion of its relevance to the Paralogisms, see Grier 2001).

¹⁸ As several commentators have noted, Kant’s cursory statement of the Paralogisms is plausibly associated with the intention to capture extremely familiar forms of argument. (Plausibly indeed, given Kant’s association of the Paralogisms with inevitable forms of Transcendental Illusion, Kant takes the arguments to be familiar from the structure of reason itself). But one central claim of this paper is that Kant’s ambitious presentation of the Paralogisms has blinded interpreters from the specific rationalist arguments they target.

A whole out of many simple substances that are thinking, thinks first when all the thoughts of each simple substance are unified in it. If each of 100 persons knew a verse from Virgil by heart, would they therefore know all of Virgil by heart?¹⁹

To appreciate Kant's vivid example, consider my entertaining a verse. There seems to be a certain unity to my consideration of the various lines. This can seem to imply that consciousness must in an important sense be simple or unified. Otherwise, if my consciousness were divided into parts, should not each part entertain only some part of the verse, while the whole of the verse is lost in the disparity of its lines? As Kant likewise writes at A352:

Representations [...] distributed among different beings, never make up a whole thought [...] and it is therefore impossible that a thought should inhere in what is essentially composite²⁰

Kant characterizes a paralogism as an argument both the major and minor premises of which are "correct", but which is fallacious in virtue of its form, since the "middle term in both premises is taken in different meanings" (R 5552, 18: 218-219). How does this formal characterization apply to the representation of the Achilles in the Paralogisms? Kant writes the following (B410-411),

The whole procedure of rational psychology is determined by a paralogism, which is exhibited in the following syllogism:

That which cannot be thought of otherwise than as subject does not exist otherwise than as subject, and is therefore substance.

A thinking being, considered merely as such, cannot be thought otherwise than as subject.

Therefore it exists only as subject, that is, a substance.

In the major premise we speak of a being that can be thought in general, in every relation, and therefore also as it may be given in intuition. But in the minor premise we speak of it only in so far as it regards itself, as subject, simply in relation to thought and the unity of consciousness, and not as likewise in relation to the intuitions through which it is given as object to thought. Thus the conclusion is arrived at fallaciously, *per sophisma figura dictionis*.

Kant's presentation focuses on the argument advanced in the First Paralogism, accusing the argument of a *sophisma figura dictionis*, i.e. a fallacy of equivocation. To see where the fallacy is supposed to lie, we should consider in some more detail the argument targeted in the

¹⁹ For an extended development of this type of argument, see also ML₁ 28: 266.

²⁰ Compare also Kant's comment that "a composite as such simply cannot think" (MVi 29:1034).

First Paralogism. Take Kant's representation of the argument in his metaphysical lectures ML₁, 28:266 (italics original):

The *I* means the subject so far as it is not predicate of another thing. What is not predicate of another thing is substance. The *I* is *the general subject* of all predicates, of all thinking, of all actions, of all possible judgments that we can pass of ourselves as a thinking being. I can only say: I am, I think, I act. Thus it is not at all feasible that the *I* would be a predicate of something else. [...] Consequently, the *I*, or the soul through which the *I* is expressed, is a substance.

We can capture the argument advanced here as grounded in a distinction between merely accidental and essential ways in which a representation can occupy the subject position in a judgment. While representations of (e.g.) actions and thoughts can accidentally serve as the subject of a judgment, they also admit of a predicative role as accidents and determinations of a thinker or agent. By contrast, the “*I*” is essentially a subject: it grounds determinations, but cannot itself be predicated of a further substance. The argument concludes that the self must be an irreducible existent, i.e. a substance.

How is this argument a paralogism? Given the definition of a paralogism, it follows that the premises of the argument are “correct.” That is, Kant must accept both the major premise asserting that if an item is essentially a subject then it must exist as a substance, and the minor premise discharging the antecedent by claiming that the self only exists as subject. Instead, Kant's criticism is that the argument *equivocates* on the notion of a subject. The major premise treats the subject as “a being in general”, i.e. at least potentially an *object*: something that, in Kant's view, must be amenable to be given in intuition. But the minor premise is grounded only in the *subjective* unity of thinking: it is the subject as she finds herself in self-conscious thought.²¹ Therefore, even as the premises of the paralogical argument are correct, the conclusion is invalid since the minor premise fails to discharge the major premise.

²¹ As Kant also characterizes the fallacy behind the paralogism in a *Reflexion*, both the minor and major premises are “correct” but the argument is “false in virtue of its form”, since the “middle term in both premises is taken in different meanings” (R 5552 18: 218-9).

However, stated in this way, Kant's critique can seem puzzling. What is wrong with the subject, as she encounters herself in relation to the unity of thought, taking herself as a "being in general"? Perhaps it is true that what the subject encounters in consciousness is not inherently a subject, but in that case the minor premise would be *false*—the argument would simply be unsound, rather than a paralogism.²² But if it is correct that the thinking self must be a subject, then why is the rationalist wrong to infer the existence of this subject as "a being in general"?

Here Kant adds a helpful note (B411, note *α*):

'Thought' is taken in the two premises in totally different senses: in the major premise, as relating to an object in general and therefore to an object as it may be given in intuition; in the minor premise, only as it consists in relation to self-consciousness. In the latter sense, no object whatsoever is being thought; all that is being represented is simply the relation to self as subject (as the form of thought). In the former premise we are speaking of *things* which cannot be thought otherwise than as subjects; but in the latter premise we speak not of *things* but of thought (abstraction being made from all objects) in which the 'I' always serves as the subject of consciousness. The conclusion cannot, therefore, be 'I cannot exist otherwise than as subject', but merely, 'In thinking my existence, I cannot employ myself save as subject of the judgment [therein involved]' This is an identical proposition, and casts no light whatsoever upon the mode of my existence.

This passage contains significant claims that shed light on the foregoing difficulties. It is not that Kant thinks there are merely insufficient grounds to posit "a being in general" from encountering the subject in thought. Rather, it is positively that "no object whatsoever is being thought" when the subject of thought is considered; rather this is merely to attend to the "form of thought" in "abstraction [...] from all objects", which is such that "'I' always serves as the subject of consciousness." So the observation in the minor premise is merely that the "I" in thinking can formally be employed only as subject, a use of "subject" entirely different than the "subject" of the major premise, which is a thing or existent in general.

Unlike the A-edition's four distinct Paralogisms, the B-edition contains merely the unified discussion printed above. So how is the rationalist's general *sophisma figura dictionis* supposed to bear on the simplicity established in the Second Paralogism? The answer is that (on

²² Note, relatedly, that the same point applies to Kant's rejection of the "verse argument". Kant does not deny that thought possesses a unity that requires a single thinking subject, lest the argument is a paralogism. Rather, the complaint is merely that the thinking subject is not an *objective* entity. This observation will structure the modest reading of Kant's criticism of rationalism provided below.

Kant's diagnosis) simplicity in the relevant sense is a spatial, experiential concept. As Kant writes, simplicity is the claim that the soul "does not consist of many subjects in space (R 18:299)."²³ But the application of this concept to the soul trades on the conception of the self as an object of intuition, which Kant has shown is not found in self-conscious reflection. Equally it follows that incorporeality and immortality cannot be established for the self, since these properties are established through inferences from simplicity.

3. *A Dilemma for Functionalists*

Given the tenor of the Paralogisms, there is a broad interpretative consensus around the claim that Kant rules out a substantial conception of the self, and more specifically, arguments for such a conception along rationalist lines.²⁴ In particular, broadly functionalist interpretations (e.g. Kitcher 1990, 2011; Dyck 2011, 2014; Bird 2006; Brook 1994) take Kant to suggest that there is nothing to the self apart from relations between mental states.²⁵ On this view, Kant sees the self (and, potentially, its simplicity) as a merely formal property of the cognitive process. These interpretations rely on Kant's apparent comments in this direction, such as (A341/B399), "the notion of the self serves only to introduce all thinking as belonging to consciousness" or Kant's claim that the self (A398) "is only the formal condition, namely, the logical unity of every thought, in which I abstract from every object." As Graham Bird characterizes the Paralogisms, for example (Bird 2006, p. 631):

The primary objection to the paralogistic argument rests on the belief that merely from the logical or grammatical point that a logical subject of judging cannot be a predicate, we cannot infer that such a subject is a real substance.

²³ On this definition the self is of course simple, but because it is not a "thing" at all: not because it is a *simple thing*. Indeed, Kant in fact rejects that the self has any actual location in space, describing its association with a body as involving the "analogue of location, but not its place" (ML₁, 28:225). See also MMr 29: 879.

²⁴ The latter claim is more precise: it is in principle consistent to take Kant to accept a substantial account of the self, while rejecting rationalist arguments for this conclusion.

²⁵ The approach seems true of Longuenesse as well, who writes: "the Critique warns us [in the Paralogisms chapter] not to consider the Gemüt or mind, the whole of our representational capacities, as a substance" (1998, pp. 7-8). Cited at Boyle Forthcoming, n. 23.

As Bird articulates the common functionalist reading, it is the existence of a substantial self that the paralogistic argument illicitly derives from the merely logical role the “I” plays in thought. On Patricia Kitcher’s specific development of the functionalist view, Kant’s conception of the positive self is best rendered as an “I”-rule, which provides a form of “synthesis” of the manifold of cognition, just as other rules do. As Patricia Kitcher has expressed the view fueled by these passages (Kitcher 2013 p. 6):

[Kant] does not have a metaphysically loaded account of a unified consciousness. He takes the ‘unity of self-consciousness’ to indicate only relations of necessary connection across mental states.²⁶

Accordingly, Kitcher’s explication of Kant’s argument in the Paralogisms is that the rationalist mistakes the “I” for an independent representation, i.e. a representation of a substance, where in fact it is merely a condition on cognition, i.e. can only operate in the presence of cognition.²⁷ In particular, on this interpretation we can represent “I” only when abstracting from its role in ongoing object cognition. Accordingly, the “I” associated with the unity of the subject is a wholly “empty” representation (A341-2/B400, A345/B406, A346/B404).²⁸

Proponents of the broadly functionalist are aware that at the very least until close to publication of the *Critique* Kant seems to accept an account of the self as a simple, noumenal substance. Critical here are Kant’s pre-critical metaphysical lectures, and especially ML₁, which are lecture notes thought to stem from the mid-to-late 1770s. As illustrated e.g. by passages like ML₁ 28:266 and MH 28:44 Kant’s metaphysical lectures seem strongly rationalist in bent, and

²⁶ See also Kitcher 1993, 2011, and for a sophisticated development of the view Brook 1994.

²⁷ E.g. Kitcher 2011, p. 194; Kitcher 2013, p. 9.

²⁸ Andrew Brook (1994) has a similar view, on which both the presence and the unity of the self are conditions on the possibility of representation of the subject, instead of amounting to awareness of a substance and its attributes (p. 171, p. 173):

(2) From the point of view of what it would be like to be a subject of representation, we must picture the subject a being and cannot picture it as a plurality of any sort

(3) If I am aware of myself as subject, I must appear to my self to be one and cannot picture myself as a plurality of any sort

portray arguments close to those in the First and Second Paralogisms. For example, consider ML₁ 28: 225-6:

The concept of I expresses: I. substantiality.—Substance is the first subject of all inhering accidents. But this I is an absolute subject, to which all accidents and predicates can belong, and which cannot at all be a predicate of another thing. Thus the I expresses the *substantiale*; for that substrate in which all accidents inhere is the *substantiale*. This is the only case where we can immediately intuit the substance.²⁹

The voluminous nature of these texts, as well as the temporal proximity of ML₁ to the *Critique*, imposes a burden on functionalist readers to explain their suggestion of a sharp break between Kant's pre-critical and critical thought. In this way, Patricia Kitcher (2011, pp. 186-188) has argued that ML₁ is not representative of Kant's later views, since Kant's argument in the Paralogisms must be understood as relying on the Transcendental Deduction, which Kitcher argues did not figure in Kant's thinking in the mid-1770s (as allegedly exemplified in Kant's notes of the Duisburg Nachlaß of 1775).³⁰

But Julian Wuerth (2010a; 2010b; 2014) has provided an impressive array of textual evidence to suggest that the dismissal of ML₁ and similar texts is unjustified.³¹ In agreement with Wuerth, I hold

(1) The Continuity of the Paralogisms: Appearances in the Paralogisms section notwithstanding, Kant's critical philosophy is consistent with his late pre-critical work in holding to a conception of the self as a simple, noumenal substance.

First, Wuerth shows that it is not true that the themes that Kitcher suggests mark Kant's departure from ML₁—in particular relevant versions of the Transcendental Deduction—are

²⁹ I return to Kant's mention of "immediate intuition" in §5 below.

³⁰ This approach is represented also by Dyck, who, despite noting "little [initial] discontinuity" between the pre-critical and critical views, suggests we should nevertheless appreciate "the dramatic changes" in Kant's view in the metaphysics of the subject, to the effect that the critical Kant "now advances the *I think* considered apart from any claim of existence as the only viable basis for the proper subject of rational psychology" (2014, pp. 183-184).

³¹ Wuerth largely focuses on providing textual evidence against the functionalist view, but at one point also emphasizes the worry that functionalist account are insufficiently capable of accounting for personal identity, insofar as they seem to embrace Parfitian account on which appropriate relations between mental states between different individuals should seem to be capable of allowing for existence of a single subject. In his discussion, Karl Ameriks (2000, pp. 261-2) sees this problem, but by Wuerth's lights Ameriks fails to take sufficiently seriously that the possibility of Kant accepting a substantial soul "can still be a natural solution" (Wuerth 2014, Chapter 2, fn. 7).

exclusive to Kant's critical period, and instead occur also in documents that plausibly precede ML₁.³² Second, Wuerth shows that the rationalist themes from ML₁ occur both in the *Critique* itself, and in the metaphysical lectures from long after the critique, such as MD, MMr and MV_i.³³ For example, in a note taken from Kant's own copy of the *Critique* (and hence obviously post-critical), Kant notes the noumenal character of the self as substance: "the I is noumenon; I as intelligence" (R CV, 23:34). Just so: "The soul in transcendental apperception is noumenal substance" (R6001 18:420-1). And as Kant writes in the post-critical *Metaphysik Mongrovius*: "If we leave aside all accidents, then substance remains, this is the pure subject in which everything inheres or is the substantial" (MMr 179 29: 771).³⁴ Indeed, the *Critique* itself grants that (A350–1, A359):

one can quite well allow the proposition *The soul is substance* to be valid.

[I]t is permitted to me to say, 'I am a simple substance', i.e. a substance the representation of which never contains a synthesis of the manifold.

The evidence of Kant's substantial conception of the self engenders a dilemma. How can we take Kant's argument in the Paralogisms seriously if Kant elsewhere accepts a positive substantial conception of the self? How can we take Kant's positive substantial conception of the self seriously if he dismisses the rationalist arguments for it in the Paralogisms?

³² See Wuerth chapter 3, section 8.

³³ For just some examples of the enormous array of critical and post-critical textual support Wuerth (e.g. 2014, p. 21) adduces: A350-1; A356; A365-6; A399-403; Prol 4: 334-5; R6001; MMr, 29: 770-1, 772, 904, 912; MvS, 28: 551; MFNS 4: 542-3; VSGE, 2: 359; ML₂ 28: 590-2; MD 28: 681-6; MK₂ 28: 754-6 759; MV_i 29: 10125-7, 1032.

³⁴ Kitcher is aware of this evidence, but her commitment to a purely functionalist account compels her to deny it expresses Kant's view. Referring to (MMr 179 29: 771), Kitcher writes (Kitcher 2011, Chapter 11, fn. 3),

The 'Metaphysics Mongrovius' asserts exactly what [Kant] is denying in his published work [...] When presumed to be accurate and read on its own, this passage suggests that Kant believed the cognizers have some cognizance of a substantial 'I.' In the published work, he is clear that this hope of Rationalist metaphysics must remain unfulfilled.

But Kitcher's attempt to isolate (MMr 179 29: 771) as unpublished material lacks plausibility given the sheer volume of Kant's remarks concerning a substantial conception of the self. It is more satisfying to cohere these remarks with the Paralogisms than to disregard them.

4. Kant's Critique of Rational Psychology

In this section I resolve the above dilemma by defending:

- (2) The Specificity of the Paralogisms:** Kant's criticisms in the Paralogisms of the Achilles are targeted at specific, identifiable versions of such arguments advanced by Kant's rationalist predecessors, as opposed to arguments of the Achilles variety *simpliciter*.

If (2) is true it promises to resolve the tension between Kant's arguments in the Paralogisms and Kant's substantial conception of the self. The Paralogisms serve to diagnose *particular* (if profound) errors in the Achilles arguments proposed by Kant's rationalist predecessors, specifically Baumgarten and Wolff. This renders Kant's concern as *the form* of an Achilles argument rather than the argument *per se*. This is the claim I defend.

As has recently been noted by several commentators, the arguments sketched in the Paralogisms bear a strong resemblance to lines of argument advanced by German rationalists.³⁵ Consider, for example, Steven Tester's recent focus on Alexander Baumgarten's *Metaphysics*—the textbook of Kant's metaphysics lectures—as a target of Kant's thinking in the First Paralogism (Tester 2014, p. 32). Like the argument in the First Paralogism, Baumgarten explicitly proceeds from the principle that each accident must have sufficient grounds in a substance since any accident or determination must have a ground, which cannot ultimately in turn be an accident or determination, at pains of regress (Me§22). Baumgarten further fits Kant's target of a rational psychologist in further inferring from the conclusion that the self must be a substance to various other attributes, including freedom (Me§755), personality (Me§641, §754), simplicity (Me§744) and immateriality (Me§757). If Kant's First Paralogism plausibly represents Baumgarten's thinking, Corey Dyck in his recent *Kant and Rational Psychology* (2014) has provided a plausible target of Kant's Second Paralogism in Christian Wolff. Wolff's argument for the existence of a simple soul strikingly recalls Kant's representation of the “verse argument.” Wolff argues that the activity of thinking must be a matter of motion (DM 738), and that the motion of different bodies

³⁵ Important here are Dyck 2011, 2014, Tester 2014.

could not explain the unified character of this particular motion (DM 739-742). Thinking must, consequently, be done by a unified subject.³⁶

For presentational purposes I rely here on Tester's and Dyck's detailed discussions of Wolff and Baumgarten, and nothing in my cursory presentation can establish Wolff and Baumgarten as the main target of Kant's First and Second Paralogisms (much less the exclusive target).³⁷ But reflecting on Kant's responses to the specifics of Baumgarten's and Wolff's arguments helps locate the upshot of the Paralogisms as far more targeted than assumed by functionalist readers.

First consider Kant's diagnosis of Baumgarten's argument as an instance of Transcendental Illusion.³⁸ This concerns Baumgarten's employment of the principle of sufficient reason. Kant agrees with Baumgarten that it is natural to infer from the fact that the self does not seem to admit of use as a predicate, to the claim that the self must be a substance. Kant makes clear in the *Prolegomena* that even as ultimately this argument rests on a confusion, its motivating idea is natural (Prol. 4:334):

Now it does appear as if we have something substantial in the consciousness of ourselves (i.e., in the thinking subject), and indeed have it in an immediate intuition; for all the predicates of inner sense are referred to the *I* as subject, and this *I* cannot again be thought as the predicate of some other subject. It therefore appears that in this case completeness in the referring of the given concepts to a subject as predicates is not a mere idea, but that the object, namely, the *absolute subject* itself, is given in experience. But

³⁶ For discussion, see Kitcher 2011, p. 48. As Richard Blackwell has made clear (Blackwell 1961, p. 346) Wolff's argument has elements that are closer to a Proclean than a Plotinian Achilles, focusing specifically on the activity of reflection:

Wolff indicates that one might argue that perception could occur in matter by means of the various parts of matter producing and sustaining the proper motions. However, there is no way in which apperception could occur in a material body. Hence the consequence is that since the soul does exercise cognition, it must be immaterial. [...] From the immateriality of the soul Wolff then concludes that it lacks part, extension, shape etc. In short, the soul is an immaterial simple substance.

In this sense Wolff's argument approaches the character of the Achilles I attribute to Kant in §5 below. However, as I show below Wolff's argument nevertheless retains the properties that place it in Kant's target area: (i) an empirical epistemology, (ii) a conclusion asserting empirical substantiality and (iii) a commitment to the claims of a broad Achilles.

³⁷ Nor is that my claim, which would contradict the clearly broad application of Kant's arguments.

³⁸ For detailed discussion of Kant's notion of a Transcendental Illusion see Michelle Grier 2001. See also Grier 1993.

this expectation is disappointed.

So why must the expectation of finding an absolute subject in experience be “disappointed”, as Kant suggests?

As Tester argues, Kant suggests Baumgarten’s employment of the principle of sufficient reason constitutes a common conflation of a respectable “logical” use of the principle on which reason seeks to (A 307/B 364) “[F]ind the unconditioned for conditioned cognitions of the understanding, with which its unity will be completed”, with a misguided “metaphysical” use of the principle that assumes (A 307–8/B 364) “when the conditioned is given, then so is the whole series of conditions subordinated one to another, which is itself unconditioned, also given.” That is, when confronted with the nature of thinking, for Kant it is natural to seek the substantial grounds of this activity. As Kant writes (Prolog 4:333)

Pure reason demands that for each predicate of a thing we should seek its appropriate subject, but that for this subject, which is in turn necessarily only a predicate, we should seek its subject again, and so forth to infinity (or as far as we get).

However, “through misunderstanding” this legitimate principle has (A309/ B366):

been taken for a transcendental principle of reason, which overhastily postulates such an unlimited completeness in the series of conditions in the objects themselves.

Kant’s last comment here is telling. The mistake in Baumgarten’s illusion-driven application of the principle of sufficient reason is that it seeks the ultimate grounds *in the objects given in experience*. The Baumgartian Achilles rests on a natural if misguided *objectification* of the grounds of our thinking: the fact that reason naturally inquires after these grounds does not mean that when some accident of the self is given in experience, so is the self as its source.³⁹ As Kant

³⁹ Especially relevant here is Kant’s notion of phenomenal substance or *substantia phaenomenon* (for discussion, see Wuerth 2014, pp. 97-100). For Kant, the genuine grounds of all accidents must be a noumenal substance, but the mind is inclined to analogize a similar relation between accidents and substances in appearance, thus giving rise to *substantia phaenomenon* (ML₁, 28: 1523; R4675, 17: 648; R4494: 17: 572). This notion of phenomenal substance plays a critical role for Kant’s account of object cognition, since its perdurant character allows for the cognition of objects in time (B226-7). Accordingly, the Transcendental Illusion holding sway over Baumgarten’s thinking is to confuse what is properly noumenal substance with *substantia phaenomenon*. As Steven Tester puts the point (2014, p. 51):

explicitly notes (Prol. 4: 333; italics mine),

human understanding is not to be blamed for its inability to know the substantiale of things [...] but rather for demanding to cognize it determinately as though it were a *given object*.⁴⁰

Turn now to Kant's specific critique of Wolff's rational psychology. As Dyck's discussion emphasizes, Wolff's rationalism is distinctive in being partially grounded in empirically-informed premises.⁴¹ Dyck notes that in this sense the common perception of Wolff as merely an expositor of Leibniz is false. That is, where Leibniz's rationalism seeks to ground metaphysical conclusions through reason alone, Wolff maintains a "mixed methodology" which, to be sure, can go beyond experience, but does not proceed in abstraction from it.⁴² As Wolff himself writes (AzDM §261), his approach is to "derive that which is found to be in [the soul] by means of experience."⁴³ One telling example of Wolff's partially experiential approach to rationalism is the inferential characterization of Descartes' cogito (DM §504):

If there is something in a being that can be conscious of something, that thing is a soul.
There exists in me something that can be conscious of something. Therefore a soul exists in me (I, a soul, exist).

As this passage illustrates, Wolff takes the cogito to infer from a premise concerning the existence of consciousness to the existence of a being capable of such consciousness. This raises the question how the minor premise is justified, which Wolff answers by appeal to an (DL c. 5,

[the] illusion consists in thinking that the application of the principle of sufficient reason will lead to an ultimate unconditioned ground of the attributes of thought that is an appearance to which spatial and temporal conception of substance applies. In contrast, Kant argues that the principle of sufficient reason will lead to a ground of the attributes of thought that is not an appearance but a thing in itself

⁴⁰ I come back to Kant's mention of an "inability to know".

⁴¹ As Dyck writes, "In contrast to the narrowly rationalistic approach to the soul which would proceed completely independently of experience, the rational psychology pioneered by the theorists of the German tradition relies essentially upon empirical psychology" (2014, p. 9).

⁴² Dyck 2014, p. 13.

⁴³ See Dyck 2008, Dyck 2014, pp. 5-10 for discussion.

par. 1) “indubitable experience” which, as an experience, is gained “when we attend to our sensations.”⁴⁴

The reason it is important to point to Wolff’s “mixed”, i.e. significantly empirical, version of rational psychology is that it is this particular aspect of Wolff’s approach that Kant criticizes. For example, as against Wolff Kant insists that properly understood rational psychology “is built on the single proposition *I think*” and does not contain “the least bit of anything empirical (A342/ B400).”⁴⁵ Just so (ML1, 28:262; italics mine)

In rational psychology the human soul is cognized *not from experience*, as in empirical psychology, but a priori from concepts. Here we are to investigate how much of the human soul we can cognize through reason.

Likewise, Kant frequently frames his opposition to the paralogical arguments explicitly in terms of their reliance on a quasi-empirical epistemology of the self (A354-55; emphasis removed):

This proposition [the ‘I think’], however, is not itself an experience [...] We have no right to transpose it onto a [...] concept of thinking being in general.

Indeed, close reading of Kant’s unified B-edition exposition of the Paralogisms reveals the same explicit focus on empiricism, taking the rationalist to conceive of the self as (B410-411): “a being

⁴⁴ Wolff’s understanding of the cogito as an inference grounded on empirical experience of consciousness is not idiosyncratic. The Wolffian German rationalist Georg Friedrich Meier similarly makes explicit that (Metaph. §481):

Daily experience teaches us that we think. This is so undeniable that even if one wanted to doubt it, doubting consists in thinking and, therefore, convinces us all the more we think. We are, accordingly, thinking beings. Now the same undeniable experience tells us that we consist of a number of parts, many of which we do not experience that they think [...] Consequently, distinct experience tells us that our thoughts are to be encountered in a part of us. Consequently, there is in us a thinking being, [or] a part which thinks, and therefore can think. This part exists because otherwise it could not think and one can always infer the existence of the cause from the existence of the effect. It follows that we have a soul, or a soul exists in us.

Similarly, it is this aspect of the cogito that the Piëtist philosopher Crusius criticizes Descartes for failing to make explicit:

[“I think”] should be: *I am conscious that I think, therefore I am*. If one wants, therefore, to infer our existence from the existence of our thoughts, one must at the outset establish the existence of our thoughts themselves by means of a sensation, namely, from inner sensation or consciousness.

For these passages and further discussion, see Dyck 2014, Chapter 6.

⁴⁵ For discussion see Tester 2014, p. 23.

that can be thought in general, in every relation, and therefore also as it may be given in intuition.”

Similarly, it is resistance to an empirically-grounded brand of rational psychology that informs Kant’s rejection of the cogito as framed by Wolff. Where the Wolffian cogito is an inference from empirical grounds, Kant holds that “*I am*, that is [...] not an inference” (AC 25: 14-15), “[n]or is the simplicity of myself (as soul) really inferred from the proposition [I think]” (A354-55). In this regard, Kant follows philosophers like Leibniz, Spinoza and, according to Udo Thiel, especially the French philosopher Johann Bernhard Merian.⁴⁶ On the approach represented by these philosophers, thought does not provide empirical grounds for an inference to the existence of a thinker, but is itself already existence-involving by, in Thiel’s terms, “[specifying] the manner in which I exist.” As Merian writes (Ma §99; emphasis removed):

When one sets the verb alongside it, one finds that thinking means as much as to exist under a certain modification, and that the proposition I think, is the same as I exist thinking.

As Leibniz expresses the same thought (NE 411):

To say *I think therefore I am* is not really to prove existence from thought, since I think and to be thinking are one and the same, and to say I am thinking is already to say I am.⁴⁷

Thus Merian writes that the self “is immediately present to itself in the sense in which, without exception, its thoughts are” (Ma §95).⁴⁸

⁴⁶ See Thiel 1996, 1998, 2001, 2011. Kitcher and Dyck both feel compelled to suggest that Kant accepts only the “I think”, rejecting any use for “I am”. As Dyck writes (2014, p. 174):

At the root of [the paralogical argument] is transcendental illusion, which Kant also characterizes in terms of conflating the formal *I think* with the empirical *I am*, a misidentification famously expressed in the Cartesian *cogito, ergo sum*.

It has been the aim of this paper to undermine the suggestion that the “I am” must be empirical, and that Kant consequently rejects it. There is also textual evidence Kant does not reject the “I am.” As Kant writes, “When something is apprehended, it is taken up into the function of apperception. **I am. I think. Thoughts are in me**” (R4676, 17: 656; emphasis original).

⁴⁷ Compare Spinoza on the same point: “I think, therefore I am is a single proposition which is equivalent to this, I am thinking” (CV 1: 234).

⁴⁸ In what sense the “I think” of the cogito is supposed to “specify” the existence of a thinker is not intended to be clear at this point. I hope to illuminate this difficult issue in §5. Here the priority is Kant’s rejection of the distinctively empirical aspect of the Wolffian philosophy.

The upshot of raising Kant's resistance to Baumgarten's and Wolff's specific versions of rational psychology is that it sheds light on the nature of Kant's arguments in the Paralogisms.

Consider that for Kant, cognition (*Erkenntnis*) depends on the impact of objects on the senses. As Wuerth states the point (Wuerth 2014, p. 127):

In our relation to all other things [other than, for Wuerth, the self], Kant tells us, we deal only with the *effects* of these things on us, which as such are colored by our manner of actively receiving these effects, in the case of humans by imposing the pure forms of intuition on them.⁴⁹

It follows that since—*pace* the rational psychologists—the “I” is not an object of sensory-affection, we can have no cognition of the subject, much less of properties like its simplicity or its immortality. This invalidates the inferences of rational psychology. As Kant writes (MMr, 29:771) “[of the *substantiale*] we cannot make the least concept, i.e., we cognize nothing but accidents.” Similarly: “I do not cognize the real simplicity of my subject (in extent of being empirically real)” (A356). Accordingly, this is the mistake of rational psychology: it takes the “I” as an object of intuition and therefore as an object of cognition, claiming knowledge of the various properties asserted in the Achilles. Such knowledge is the “expectation [which] is disappointed” (Prol. 4:333).

5. Kant's Pure Achilles

We should now return to the Achilles as discussed in §1. As Julian Wuerth writes, Kant's central target is the *broad* Achilles: the inferences from the existence of the soul as a substance to various other attributes (Wuerth 2014, p. 116):

in attempting to make sense of Kant's argument against the rational psychologists, it should always be remembered that Kant repeatedly emphasizes that the sole purpose of the rationalists' ventures in psychology is to establish the immortality of the soul.

⁴⁹ For a book-size development of this aspect of Kant's view, see Langton 1998. Wuerth helpfully relates (chapter 3, fn. 3) the effect-dependence of Kant's cognition to Kant's adherence to the *principium generale commercii*, i.e. the thought that—as Leibniz famously denies—substances not only *do* genuinely interact, but that moreover “all influence in the world is partly the effect of the active on the passive, partly the counter effect of the latter” (R 4704, 17: 681). Adherence to this principle goes a significant way to motivating Kant's view that objects can only be known through their effects on the knowing subject.

As Kant himself explicitly notes, rationalist concern with the soul resides primarily in its supposed permanence, and especially its immortality (Prol. 4: 335):

Though we may call this thinking self (the soul) substance, as being the ultimate subject of thinking which cannot be further represented as the predicate of another thing, it remains quite empty and inconsequential if permanence—the quality which renders the concept of substances in experiences fruitful—cannot be proved of it. But permanence can never be proved of the concept of a substance as a thing in itself, but only for the purposes of experience.⁵⁰

For Kant the predicates of the broad Achilles fail because they lay claim to the status of cognition, and more specifically, because they are associated with an experiential concept of substance and depend on spatio-temporal structures applicable only to experience. As Kant writes (A356):

one by no means [...] cognizes anything about the soul that one really wants to know, for all these predicates are not valid of intuition at all, and therefore cannot have any consequence that could be applied to objects of experience; hence they are completely empty.⁵¹

As this last passage compellingly suggests, *this—pace* the functionalist reading—is the significance of Kant’s assertion (A356) that the “I think” is merely “empty”, of no “value”, and is of no “use *in concreto*.” As Kant stresses (A356):

everyone must admit that the assertion of the simple nature of the soul is of unique value only insofar as through it I distinguish this subject from all matter, and consequently except it from the perishability to which matter is always subjected. It is really only to this use that the above proposition is applied, hence it is often expressed like this: the soul is not corporeal.

In this section I will argue that it does not follow from the “emptiness” of rationalist discourse that Kant rejects a substantial account of the self (much less does the positive functionalist claim

⁵⁰ Just so Kant writes in the post-critical *Metaphysik Mongrovius* (MMr 29:912; italics mine):

one believes to be done with [the question of immortality] since substance is perdurable. But since *the soul is recognized as substance only through the I*, we do not at all know whether it is substance *in the sense that* as such it could not perish.

Note the expressions I emphasize. Kant does not deny that the soul is recognized as substance, but merely notes that this is “only through the I”. And the conclusion is specifically limited to the claim that the soul is not known to be substance “in the sense that” it is imperishable. This fits my claims in this paper.

⁵¹ This tenor of Kant’s critique extends to Mendelssohn’s Achilles argument for the immortality of the soul mentioned at the outset of this paper (the criticism is at MFNS 4:542-3).

follow that the “I” must be understood merely in terms of functional relations between mental states.)⁵² By contrast, I suggest Kant’s critique of the empirically-oriented Wolffian/Baumgartian approach makes room for a distinct, purely rationalist substance-metaphysics of the self. In particular, I will defend:

(3) Kant’s Critical Achilles: Kant’s critical thought offers a version of the Achilles argument.

Consider the following contrast Kant draws (A546/B574):

the human being, who is otherwise acquainted with the whole of nature solely through sense, is acquainted with himself also through pure apperception.

Consider similarly (SGE 20: 359):

[T]he simple is not given in outer intuition. In the inner there is the simple but only in the *subject* of consciousness...in so far as it *thinks* not insofar as it has an *intuition* of itself through the inner sense therefore it is also not given for *knowledge*.

In both of these passages Kant explicitly opposes cognition through sensory awareness to the subject’s relation to herself. For Kant, the self is “acquainted” with herself through “pure apperception.” Likewise, there is simplicity in the self but “but only in the *subject* of consciousness...in so far as it *thinks* not insofar as it has an *intuition* of itself.” Compare similarly (B429; B277; italics mine)

In the consciousness of myself in mere thinking I am the being itself, about which, however, nothing yet is thereby *given to me* for thinking.

of course, the representation **I am**, which expresses the consciousness that can accompany all thinking, is that which immediately includes the existence of the subject in itself, but not *yet any cognition of it, thus not empirical cognition, i.e. experience*.

⁵² Compare (A356; A350–1):

Now mere apperception (“I”) is substance in concept, simple in concept, etc., and thus all these psychological theorems are indisputably correct. Nevertheless, one by no means thereby cognizes anything about the soul that one really wants to know, for all these predicates are not valid of intuition at all, and therefore cannot have any consequence that could be applied to objects of experience; hence they are completely empty. For that concept of substance does not teach me that the soul endures for itself.

[O]ne can quite well allow the proposition *The soul is substance* to be valid, if only one admits that this concept of ours leads no further, that it cannot teach us any of the usual conclusions of the rationalistic doctrine of the soul.

Here Kant seems still stronger in stating the above contrast. In self-consciousness “I am the being itself” and such consciousness “immediately includes the existence of the subject in itself”—but these are not experiences, and therefore not cognitions.

Julian Wuerth has described Kant’s epistemology of the self as “immediatism”.⁵³ On Wuerth’s view, Kant’s consistent view throughout the pre-critical metaphysics lectures, the critical period and the post-critical lectures is that a subject has immediate access to herself as a noumenal substance *simply in virtue of being this substance*. As Wuerth characterizes his reading in an important passage (2014, p. 76; italics original):

Our relation to ourselves is the *one* instance in which we do not relate to a substance by virtue solely of its *effect* on us, but instead by virtue of *being* the substance, and, as *rational* substances, by necessarily being *conscious* of being a substance, and so of *having* powers by means of which we ground accidents. This view that we can have some form of epistemic access to *our own* noumenal selves but not to *other* noumenal substances is therefore not inconsistent, but is instead of a piece, with the guiding idea behind transcendental idealism. For in this lone instance can we be conscious of being the something that is active in addition to being conscious of any effects of this activity. Rather than reflecting an immature epistemology that fails to recognize the barriers precluding epistemic access to other things in themselves, Kant’s immediatism therefore reflects the unique place of our relation to ourselves within this system of transcendental idealism.⁵⁴

That is, where cognition of objects depends on an object’s effects on the senses, for Wuerth apperceptive awareness allows for a type of awareness of the self that is more direct (and therefore not phenomenal). Since consciousness is apperceptive, I can be aware of myself as subject simply by *being* the subject. In this regard, Wuerth’s “immediatism” has strong textual

⁵³ Wuerth 2006; 2010a; 2010b; 2014.

⁵⁴ For Wuerth, the main shift in Kant’s pre-critical and critical views—and the origins of the Paralogisms—is not the abandonment of a previous commitment to rational psychology, but rather a change in the shape of Kant’s “immediatist” epistemology. In Kant’s pre-critical metaphysics lectures, Kant conceives of the subjects access to herself as an “immediate intuition.” As Kant writes in the pre-critical metaphysics lectures (ML₁, 28: 226),

[our access to the soul] is the only case where we can immediately intuit the substance. Of no thing can we intuit the substrate and the first subject; but in myself I intuit the substance immediately⁵⁴

Kant pre-critical “immediate intuitions” are to be sure not properly experiential in not facilitating any cognition since they are wholly “indeterminate” and fail to reveal any attributes (see R4493 17: 572, according to which “the predicates are missing.”). Nevertheless, the critical Kant moves away from mentioning intuition altogether in this context in favor of an apperceptive immediatism.

support. Both in the pre-critical and critical writings Kant stresses the opposition between cognition of objects through sense and “immediate” access to the self. Nevertheless, I suggest Wuerth does not succeed in fully clearing up the philosophical questions surrounding Kant’s view. Below I suggest that attributing a version of an Achilles argument provides a more complete understanding of the distinction Wuerth offers.

There are two questions that Wuerth’s treatment does not address. First, Wuerth does not explain why for Kant there should be a form of “immediate” self-access, and why no sensory impacts are required for this type of access. Second, nothing in Wuerth’s account makes obvious why the subject accessed in apperception should be a *substance*.⁵⁵ This is where I suggest attributing to Kant a version of a bit of Achilles reasoning has explanatory value.⁵⁶

Kant’s Pure Achilles

P1: There is a constitutive unity between the subject and the object of apperception
P2: The unity at P1 is metaphysically primitive in such a way as to be constitutive of a simple, unified substance.
C1: The human self or subject is a simple unified substance

In the terms of §1, this is a narrow Proclean Achilles. The argument has two distinctive features. The Kantian Achilles fills out the “unity” or “simplicity” of P1 in terms of a constitutive identity between the subject and the object of an act of reflection. At P2 the argument moreover asserts this identity to be metaphysically primitive. I flesh out this argument below, but let me first indicate how the Achilles helps answer the questions left by Wuerth’s account. Where Wuerth’s account might seem to posit an unexplained non-sensuous epistemic relation a subject bears to herself, the Achilles explains this relation. A subject’s self-consciousness is constitutive of her being. It follows that self-awareness is “immediately” provided by being the subject. It also follows this self-awareness could not be sense-mediated, since the object and subject of self-awareness are identical and sense mediates access of one item to another. Per P2 the Achilles also

⁵⁵ To be sure, if access to the self is not through sense it is obvious the self cannot be phenomenal, and therefore might seem noumenal. But the question is why the self should be any “item” *at all*. Wuerth seems to argue simply from Kant’s comments in this area, and not seek to explain Kant’s commitment.

⁵⁶ I should explicitly note that this is an interpretative claim. I defend the merits of this view of self-consciousness in the next chapter.

explains Kant's use of substantial language about the self.

In his discussion of Kitcher's *Kant's Thinker*, Sebastian Rödl suggests that Kitcher's progress notwithstanding, her work requires further development in the area of the way self-consciousness involves a "certain consciousness of unity: a consciousness of unity which is nothing other than this unity" (Rödl 2013, p. 216). Rödl describes here a remarkable relation between (a form of) consciousness and (a form of) unity: a consciousness *of* a unity that is at the same time *itself* the unity in question. Rödl identifies this peculiar consciousness with the activity of synthesis, and the unity with synthetic unity.⁵⁷ As Rödl fills out Kant's characterization of synthesis (B130; translation Rödl added), "Combination [synthesis; S.R.] is a representation [consciousness; S.R.] of the synthetic unity of the manifold."⁵⁸ *Pace* Kitcher who argues that the identity of the "I" requires an account beyond the nature of apperceptive awareness, Rödl further identifies this synthetic unity with the identity of self-consciousness as expressed by the "I", noting that "The unity of the subject is its [reflective] act [...] Such is the unity of a thinker" (Rödl 2013, p. 216).⁵⁹

Rödl's conception of a "consciousness of unity which is nothing other than this unity" provides the type of unity that lies at the heart of Kant's version of the Achilles. A first step in appreciating the way Rödl's view might support a form of the Achilles is to consider a view of self-awareness on which it is an activity that takes the "I" both as its subject and its object. The point is illustrated by the above-discussed Merian conception of the cogito. On this conception, the reflective awareness embodied in the subject's self-consciousness of her thinking finds not merely an act of thought, but in this same act a subject of thought—the "I" of the "I think." But

⁵⁷ It follows that for Rödl synthesis is not an activity productive of unity, since "synthesis were production, it could not be a consciousness of unity, for the unity produced would be different from the synthesis that produced it" (2013 p. 216).

⁵⁸ One may take exception to Rödl's rather tendentious characterization of B130. It does not matter to my purposes.

⁵⁹ Rödl's reading—to which I am sympathetic—is in some sense in agreement with Melnick 2009, in that both take the "I" to be constitutively related to some activity. But where Rödl takes the activity to be reflective, Melnick characterizes it as a "marshaling" that produces "order from diversity". I think that Rödl's conception is required as the background against which to understand the "order" of thought, and I do not see how Melnick's suggestion of a mere creation of order can constitute a self.

this “I” is the very same “I” that is the subject of the original act of self-reflection. That is, this act of self-reflection is executed by the “I” and at the same time finds this “I” as the object of its awareness.

But pointing to an (alleged) identity of object and subject in self-consciousness is not sufficient to see how the reflective act bears a distinctive unity (much less a substantial unity). Any reflexive act bears the same subject and object. By contrast, the second step in appreciating Rödl’s view is the suggestion that the reflexivity of the self-conscious act is *constitutive* of both of its subject and its object. The “I” of reflection is constitutively related to the awareness it bears of itself. (By identity, so is the “I” that is the object of self-awareness). It is popular to say that the “I” is an expression that refers *de jure* to the subject of the locution (e.g. Peacocke 2014). But this is to focus merely on the expression and miss that the “I” expresses the reality of a version of consciousness. That is, the “I” evinces (since it presupposes) a subject that would say “I.” It is in this way that the “I” is constitutively associated with self-awareness: “I” is merely the articulation of the type of reflexive act displayed by subject’s awareness of herself. To put matters in a slogan, the “I” is essentially “double”: its nature is such that it is always at once the subject and the object of an act of awareness.

This is a substantial metaphysical view, which raises significant questions of its own. However, it is worth stressing that the conception of the self as constitutively associated with self-awareness has a rich historical heritage. A *locus classicus* is Augustine’s conception of the perpetually self-knowing mind.⁶⁰ As Ursula Coope has suggested, a similar commitment may seem to be found in the Thomistic understanding of the intellect (Aquinas CPA: preface; discussed at Coope 2013, p. 8):

Reason (*ratio*), however, is not only able to direct the act of inferior parts, but is also the director of its own act *actus sui directiva*). For the ability to reflect on itself is peculiar to

⁶⁰ *DeTrin*, 10.5.7, 14.6.9, 15.15.25. See also *DeTrin* 9.3.3, for a parallel to Kant’s conception of sense-independent apperception “Just as the mind itself gathers knowledge of corporeal things through the senses of the body, so [it gathers knowledge] of incorporeals through its own self. Therefore it also knows itself through itself.” Cited and discussed at Cory 2014, pp. 19-21.

the intellectual part. For intellect understands itself and similarly, reason can reason about its own act.⁶¹

Although the passage does not explicitly suggest a constitutive conception (merely stating that the “ability to reflect on itself is peculiar to the intellectual part”), it is plausible to understand Aquinas as endorsing the Proclean thought that rather than following from an inference to the best explanation, “the intellectual part” is constitutively associated with its self-reflective character. Finally this line is famously expressed by Sartre when he writes “The Ego is an object apprehended but also an object *constituted* by reflective knowledge.”⁶²

Still, even if the above three steps are granted, it is nevertheless a further question why the constitutively reflective “I” must be considered a substance.⁶³ Consider in this regard a broad version of the Proclean Achilles provided by Aquinas (SCG II.49):

No body’s action reflects on the agent, for it is proved in the *Physics* that no body reflects on itself except in respect of a part, so that, namely, one of its parts be mover and the other moved⁶⁴

Evidently this argument is not available to Kant since it relies on a spatial concept of part, and Kant’s self is neither part-bearing nor part-less in this sense. But the argument displays a more abstract structure which may serve to illustrate why Kant would speak of the self in substantial terms.

In a more abstract sense, reflection might at first glance seem to be a “bipartite” activity:

⁶¹ For a treatment that places Aquinas in opposition to the neoplatonist tradition, see Cory 2014, Chapter 1.

⁶² This is from Sartre’s 1937 *The Transcendence of the Ego*, translated in Sartre 2004. Discussed at Peacocke 2014, pp. 220-221. To be sure, there are complexities to Sartre’s account (such as a priority of consciousness to self-consciousness (Sartre 2004, p. 34)) that I am not here in a position to discuss. Of course, the thesis also has a rich history in German Idealism, for example in Fichte, who calls the “I” “unconditioned” and “self-grounded” (SW, I, pp. 95-6). For discussion see Neuhauser 1990, p. 43.

⁶³ One might here recall the considerable evidence so far: (i) Kant repeatedly speaks of the “I” in substantial terms; (ii) once freed from empirical associations Kant has no obvious grounds for resistance a substantial view (iii) Kant’s usage of the terms “form” and “unity” as associated with the “I” track directly the Leibnizian conception of a substance, providing salient precedent. Moreover, resistance to attributing a substantial subject to Kant may in part derive from remaining “objective” associations with the realm of substances. But for Kant substances are emphatically *not* objective (neither are the substances of Leibniz’s Idealism.)

⁶⁴ As Coope notes, Aquinas supports this inference through a neoplatonist reading of Aristotle’s conception of self-movement, discussed at (CP, VII, lecture 1.890). As Aquinas also writes: “But it is not possible that anything material should change itself, rather one is changed by another” (ST Ia 187 a3 ad3).

an activity in which one “part” reflects on another “part.” But by Proclean lights this conception does not articulate genuine reflection, but rather maintains the logical structure of one item relating to another (specifically, the reflecting state relating to the reflected state). By contrast, genuine reflection requires one item relating to *itself*. I want to suggest this is an abstract sense in which reflection is “partless” that can support the above argument in a Kantian setting. If Kant accepts, in broadest outline, the Leibnizian view that a substance is defined in terms of an “activity” with a certain “unity” the question is whether reflection can be understood in light of some non-reflective activity or must be considered *sui generis*. It must be *sui generis*: no non-reflective activity relates to itself in the way reflection does. This self-relation is constitutive of reflection. Accordingly, reflection as an activity possesses a distinct “simplicity” or “unity” that marks it as a type of substance. Since reflection is constitutive of the “I”, we can accordingly say that the subject is a substance.

6. Objections

But finding a substantial conception of the self in Kant is controversial. Here I want to discuss an especially subtle objection raised in a recent comment by Matthew Boyle on Longuenesse’s treatment of the Paralogisms. Boyle begins by noting Longuenesse’s broadly functionalist treatment of the upshot of the Paralogisms: (Boyle Forthcoming, fn. 23):

Longuenesse notes the parallel between Kant’s general remarks on substances, powers, and acts and his talk of mental acts and mental powers, but suggests that “one should approach this parallel with caution” since “the Critique warns us [in the Paralogisms chapter] not to consider the Gemüt or mind, the whole of our representational capacities, as a substance.”

Subsequently, Boyle contrasts a nuanced, partially accepting reading of Kant’s substantial language (Ibid. I have underlined a passage to which I will return below):

[But] Kant himself – according to the lecture notes of one of his students, anyway – mentions the power of thinking in a general discussion of the relation between substance, power, and act. These lectures were given in 1783, shortly after the publication of the first edition of the First Critique, and in them he repeats essentially the same criticisms of Rational Psychology that he makes in the Paralogisms chapter. It seems, then, that he must not think there is a tension between bringing the framework of substances, powers

and acts to bear on the mind in this way and his attack on the fallacious conclusions of the rational psychologists. [...] What Kant denies in the Paralogisms is that we are entitled to infer that the “I” that thinks is really a substance in itself (which would entail its permanence, immutability, etc.) from the fact that it must be conceived in terms of the logical categories characteristic of substances. He does not deny that we must conceive of our mind as having the logical nature of a substance, and thus of our mental powers as having the kind of unity the powers of a substance have. On the contrary, he affirms this, although he denies that it entitles us to draw the conclusions of the Rational Psychologists: [E]veryone must necessarily regard Himself as substance, but regard his thinking only as accidents of his existence and determinations of his state. [O]ne can quite well allow that the proposition The soul is substance holds, if one only admits that this concept [...] cannot teach us any of the usual conclusions of the doctrine of rational psychology, such as the everlasting duration of the soul through all alterations, even the person’s death – [if only one admits, in other words] that it signifies a substance only in idea but not in reality. (A349-351) The connection I am drawing between Kant’s remarks about our power of understanding and his thinking about the categories of substance, power, and act does not require any claim about the real nature of the subject to whom these powers belong; it simply brings out the consequences of regarding the mind as a substance “in idea,” as Kant says we must (cp. also A672/B700). The kind of unity our mental powers possess is logically speaking the kind of unity that the powers of a substance possess, even if we cannot conclude that this logical unity corresponds to the sort of real unity that rationalist metaphysicians sought to infer from it. This is a difficult topic, however, which deserves a fuller treatment than I can give it here.⁶⁵

As Boyle himself remarks, this footnote clearly does not amount to a full reading of the Paralogisms. Nevertheless Boyle makes an interesting suggestion, which displays significant sensitivity to the themes suggested in this paper, but also diverges from my reading in a critical respect. Boyle shows sympathy to the suggestion that understanding Kant’s conception of the relation between the “I” of apperception and its representations involves taking seriously Kant’s talk of the self as substance. But nevertheless Boyle’s view reads the Paralogisms as denying that the “I” is really a substance, and conversely Boyle holds that Kant gives a merely logical gloss on talking of the self as a substance. For Boyle, it is speaking *as though* the self were a substance that helps Kant articulate his view of apperception. In this way, Boyle’s view seems to occupy an attractive middle ground between the position defended here and the consensus functionalist view, which disavows substantial talk. In particular, Boyle’s alternative raises the question what could be gained by insisting that Kant holds a substantial conception of the self: the alternative seems to retain all the advantages, without any of the interpretative costs.

I have several brief responses to Boyle’s suggestion, which largely pick up on the

⁶⁵ Thanks to Tyke Nuñez pointing me to this passage.

underlined segment of the above-quoted passage. First, if for Kant the “I” is not a substance, then there is a question why in a “logical” sense it should be regarded as such. To be sure, this question seems answerable, but it is an unexamined feature of Boyle’s suggestion that it is plausible or even intelligible to think the self must be considered “logically” substantial if it is metaphysically *precisely* not. What undergirds the application of substantial language if it is strictly false or misleading? By contrast, if Kant’s self is substantial the answer is straightforward.

Second, there is a question concerning the motivation for Boyle’s insistence on rejecting a substantial interpretation of Kant’s view of the self.⁶⁶ Consider the underlined segment of Boyle’s above sketch of Kant’s thinking, where Boyle characterizes the “fallacious conclusions of the rational psychologists” as the claim “that we are entitled to infer that the “I” that thinks is really a substance in itself (which would entail its permanence, immutability, etc.)”. However, as I have argued in the previous sections of this paper, I think this is precisely wrong. It is not thinking of the self as a substance *in itself* which entails properties like permanence, immutability and so on. Those are (potential) properties of *empirical* substances. The mistake of philosophers like Wolff and Mendelssohn is precisely to draw from the role of the “I” in thought conclusions concerning the “I” as though it were a potential object of intuition. But for Kant to speak of an object of intuition) is emphatically not the same as speaking of a substance. Kant’s critique of the self as accessed through sensibility serves precisely to exempt and make room for access through pure apperception. So once the treatment of the self as *objective* has been rejected, it is unclear what further arguments are supposed to lead Kant to reject the self as substantial.⁶⁷

Third, there is a further level on which Boyle fails to motivate his sharp distinction between attributing to Kant a full-blown substantial conception of the self, and substantial

⁶⁶ Clearly some motivation is required, on account that this interpretation requires in some sense disavowing Kant’s apparently straightforward declarations that the subject *can* be considered a substance (provided that suitable errors are avoided).

⁶⁷ Indeed, Kant explicitly notes that the sense of “substance” at issue in the Paralogisms is the “empirically usable concept of a substance” (A349), since substance as it figures in pure thought cannot be found to ground the properties the rationalist is interested in.

language that functions merely “as if”. My suggestion is that Boyle’s distinction depends on an inflated and inappropriately demanding conception of “substance.” On the Leibnizian conception that served as the background for much of Kant’s thinking, the foundational feature of a substance is its possession of a “true unity”, a “form”, all conceived on the model of “what is called ‘me’” (AG 89). These are all features that characterize Kant’s view of apperceptive consciousness, which moreover admits of characterization in substantial terms (as Boyle is right to acknowledge). To be sure, Boyle presumably rejects Leibnizian substance talk, but it is important to spell out on which grounds.⁶⁸ What is to stop Kant from thinking of apperception in actual substantial terms, as opposed to merely “as if”? This question is especially salient in light of Kant’s specification of a *different* notion of substance as distinctly empirical (A349). The fundamental point is that despite Boyle’s appreciating the relevance of substantial talk to Kant’s view of self-consciousness, he retains the functionalist commitment to sharply distinguishing Kant’s talk of a “logical” and a “substantial” characterization of the subject and its unity in thought. But if I have been correct in this paper, this dichotomy is false: the “logical” unity of the subject can ground and indeed *be* its substantial unity.⁶⁹

References

KANT references to *Critique of Pure Reason* (1997), Guyer, P. and Wood, A.W. Cambridge, UK: Cambridge University Press, standard A/B notation. Otherwise:

AA = (1902-) *Kants gesammelte Schriften*, edited by the Königlich Preussischen (now Deutschen) Akademie der Wissenschaften, G. Reimer [now Berlin: De Gruyter].

⁶⁸ Indeed, there are significant independent grounds for thinking that Kant does *not* fundamentally abandon the presuppositions of the Leibnizian philosophy. For powerful arguments for this thought see e.g. Anja Jauernig (2008).

⁶⁹ An admitted remaining source of discomfort is Boyle’s citation of Kant’s comment that (A351) the “I” “signifies a substance only in idea but not in reality.” There are two responses, however. First it is to be granted that Kant’s consuming focus in the Paralogisms is the negative case against the substantial self of rational psychology: the confinement the “I” to the realm of ideas may seem a rhetorical flourish in this regard, not to be taken to rule out any substantial self in a sense different from that of rational psychology. More satisfyingly, however, I submit that Kant’s notions of “idea” and “reality” are sufficiently complicated to make room for my reading, as they would if “reality” could be understood in empirical terms, and “idea” in terms of self-consciousness. But clearly this passage requires further explication than I am giving it here.

AC = *Anthropology Collins* (AA 25).

Anthr. = (2006) *Anthropology from a Pragmatic Point of View*. Trans. Louden, R. Cambridge, UK: Cambridge University Press.

ML₁ = (1997) *Metaphysik L₁* (AA 28). In *Lectures on Metaphysics*. Trans. Ameriks, K. and Naragon, S. Cambridge University Press.

ML₂ = (1997) *Metaphysik L₂* (AA 28). In *Lectures on Metaphysics*. Trans. Ameriks, K. and Naragon, S. Cambridge University Press.

MMr = (1997) *Metaphysik Mongrovius* (AA 29). In *Lectures on Metaphysics*. Trans. Ameriks, Karl and Naragon, S. Cambridge, UK: Cambridge University Press.

MH = (1997) *Metaphysik Herder* (AA 28:39 – 53). In *Lectures on Metaphysics*. Trans. Ameriks, K. and Naragon, S. Cambridge, UK: Cambridge University Press.

MD = (1997) *Metaphysik Dona* (AA 28) in *Lecture on Metaphysics*. Transl. Ameriks, K. and Naragon, S. Cambridge, UK: Cambridge University Press.

MK₂ = (1997) *Metaphysik K₂* (AA 28). In *Lecture on Metaphysics*. Trans. Ameriks, K. and Naragon, S. Cambridge, UK: Cambridge University Press.

MvS = *Metaphysik v. Schön* (AA 28).

MVi = *Metaphysik Vigilantius* (AA 29). In *Lecture on Metaphysics*. Trans. Ameriks, K. and Naragon, S. Cambridge, UK: Cambridge University Press.

MFNS = (1970) *Metaphysical Foundations of Natural Science* (AA 4). Trans. Ellington, J.W. Indianapolis: Hackett Publishing.

ProL. = (2001) *Prolegomena for Any Future Metaphysics*. Transl. Ellington, J.W. Indianapolis: Hackett.

R = *Reflexionen*, Kants handschriftlicher Nachlaß, (AA 14-19).

R CV = *Reflexionen* in Kant's copy of the *Critique of Pure Reason* (AA 23).

VSGE = *Vorarbeiten zur Schrift gegen Eberhard* (AA 20).

SGE = (1973) *Schrift Gegen Eberhard* (AA20). In *The Kant-Eberhard Controversy*. Trans and eds. Allison H. E. Baltimore: John Hopkins University Press.

LEIBNIZ:

AG = (1989) *Philosophical Essays*. Eds. & trans. Garber, D. & Ariew, R. Indianapolis: Hackett.

NE = (1996) *New Essays on Human Understanding*. Eds. and trans. Bennett, J. and Remnant, P. Indianapolis: Hackett.

WOLFF:

DM = (1983) *Deutsche Metaphysik* or *Der vernünftigen Gedancken von Gott, der Welt und der Seele des Menschen, auch allen Dingen überhaupt* ([Halle], 1751); repr. in Wolff, *Gesammelte Werke*, division 1, vol. ii, Georg Olms. (Trans. at Dyck, C. 2014).

DL = (1965) *Deutsche Logik* or *Vernünfftige Gedanken von den Kräften des menschlichen Verstandes und Ihrem richtigen Gebrauche in Erkänntnis der Wahrheit* repr. in Wolff, *Gesammelte Werke*, division 1, vol. i, Georg Olms. (Trans. at Dyck, C. 2014).

AzDM = (1983) *Deutsche Metaphysik (Additional Comments)* (*Der vernünftigen Gedancken von Gott, der Welt und der Seele des Menschen, auch allen Dingen überhaupt, anderer Theil, bestehend in ausführlichen Anmerckungen* ([Frankfurt: 1740]); repr. in Wolff, *Gesammelte Werke*, division 1, vol. iii, Georg Olms. (Trans. at Dyck, C. 2014).

BAUMGARTEN:

Me = (2013) *Metaphysics*. Trans. Fugate, Cortuney D. London: Bloomsbury Publishing.

FICHTE:

SW = (1762 – 1814) *Johann Gottlieb Fichtes sämtliche Werke*. Berlin: de Gruyter.

SPINOZA:

CV = *Collected Works*.

MEIER:

Metaph. = (2007) *Metaphysik*, 2nd edn. 4 vols (1765), repr. in Christian Wolff, *Gesammelte Werke*, Pt. 3, Vols. 108.1-108.4, Georg Olms (Dyck, C., 2014 trans.).

MERIAN:

Ma = (1751) ‘Memoire sur l’Apperception de sa Propre Existence’. In *Histoire de l’Academie Royale des Sciences et Belles-Lettres* (1749). Berlin.

AQUINAS:

ST = (1949) *Summa Theologica*. Transl. Fathers of the Dominican Province. New York: Benzinger Brothers.

SCG = (2014) *Summa Contra Gentiles*. London: Aeterna Press.

CPA = (2008) *Commentary on the Posterior Analytics*. Transl. Bernquist, R. South Bend, ID: Dumb Ox Books.

CP = (1999) *Commentary on Aristotle’s Physics*. Trans. Blackwell, R. J., Spath, R. J., Thirkel, W. E. South Bend: Dumb Ox Books.

AUGUSTINE:

DeTrin = (2002) *On the Trinity*. Ed. Matthews, G.B. Cambridge, UK: Cambridge University Press.

PLOTINUS:

EA = (1992) *Enneads*. Trans. MacKenna, S. New York: Paul Brunton Philosophic Foundation.

PROCLUS:

ET = (1963) *The Elements of Theology*. Ed. Dodds, E.R. Oxford: Clarendon Oxford Press.

CT = (2007) *Commentary on the Timeaus*. Ed. Tarrant, H. Cambridge, UK: Cambridge University Press.

Further References

- Ameriks, K. (2000) *Kant and the Fate of Autonomy*. Cambridge, UK: Cambridge University Press.
- Barnett, D. (2008) 'The Simplicity Intuition and Its Hidden Influence Influence on the Philosophy of Mind'. *NOÛS* 42: 308-355.
- Barnett, D. (2010) 'You are Simple'. In *The Waning of Materialism*. Ed. Koons, R. Oxford: Oxford University Press.
- Blackwell, R. (1961) 'Christian Wolff's Doctrine of the Soul'. *Journal of the History of Ideas* 22 (3): 339-354.
- Bird, G. (2006) *The Revolutionary Kant*. La Salle: Open Court Publishing.
- Boyle, M. (Forthcoming) 'Kant on Logic and Laws of the Understanding'. In *Logical Aliens*. Ed Travis, C. Cambridge, MA: Harvard University Press.
- Brook, A. (1994) *Kant and the Mind*. Cambridge, UK: Cambridge University Press.
- Burge, T. (1996) 'Our Entitlement to Self-Knowledge'. *Proceedings of the Aristotelian Society* 96: 91-116.
- Coope, U. (2013) 'Aquinas on Judgment and the Active Power of Reason'. *Philosophers' Imprint* 13 (20): 1-19.
- Cory, Therese Scarpelli (2014) *Aquinas on Human Self-Knowledge*. Cambridge, UK: Cambridge University Press.
- Dyck, C. (2008) 'The Subjective Deduction and the Search for a Fundamental Force.' *Kant-Studien* 99 (2): 152-179.
- Dyck, C. (2011) 'A Wolff in Kant's Clothing: Christian Wolff's Influence on Kant's Accounts of Consciousness, Self-Consciousness and Psychology' *Philosophy Compass* 6 (1): 44-53.
- Dyck, C. (2014) *Kant and Rational Psychology*. Oxford: Oxford University Press.
- Engstrom, S. (2013) 'The Unity of Apperception.' *Studi Kanti* 26: 37-54.
- Grier, M. (1993) 'Illusion and Fallacy in Kant's First Paralogism.' *Kant-Studien* 84 (3): 257-282.
- Grier, M. (2001) *Kant's Doctrine of Transcendental Illusion*. Cambridge, UK: Cambridge University Press.
- Harper, W. (2008) 'Kant on the Achilles Argument.' In *The Achilles of Rationalist Psychology*. Eds. Lennon, T. and Stainton, R. Dordrecht: Springer.
- Henry, D. (2008) 'The Neoplatonic Achilles'. In *The Achilles of Rationalist Psychology*. Eds. Lennon, T. and Stainton, R. Dordrecht: Springer.

- Jauernig, A. (2008) 'Kant's Critique of the Leibnizian Philosophy: *Contra* the Leibnizians, but *Pro* Leibniz'. In *Kant and the Early Moderns*. Eds. Garber, D. and Longuenesse, B. Princeton: Princeton University Press.
- Keller, P. (1998) *Kant and the Demands of Self-Consciousness*. Cambridge, UK: Cambridge.
- Kitcher, P. (1990) *Kant's Transcendental Psychology*. Oxford: Oxford University Press.
- Kitcher, P. (2011) *Kant's Thinker*. Oxford: Oxford University Press.
- Kitcher, P. (2013) 'Précis.' *Philosophy and Phenomenological Research* 87 (1): 200-212.
- Kulstad, M. (1982) 'Some Difficulties in Leibniz's Account of Perception.' In *Leibniz: Critical and Interpretative Essays*. Ed. Hooker, M. Minneapolis: Minnesota Archive Editions.
- Langton, R. (1998) *Kantian Humility*. Oxford: Oxford University Press.
- Lennon, T. and Stainton, R. (2008) 'Introduction'. In *The Achilles of Rationalist Psychology*. Eds. Lennon, T. and Stainton, R. Dordrecht: Springer.
- McRae, R. (1976) *Perception, Apperception and Thought*. Toronto: Toronto University Press.
- Melnick, A. (2009) *Kant's Theory of the Self*. London: Routledge.
- Mendelssohn, Moses (2007) *Phädon or On the Immortality of the Soul*. Noble, P. New York: Peter Lang.
- Messina, J. (2014) 'Kant on the Unity of Space and the Synthetic Unity of Apperception'. *Kant-Studien* 105 (1): 5-40.
- Neuhouser, F. (1990) *Fichte's Theory of Subjectivity*. Cambridge, UK: Cambridge University Press.
- Proops, I. (2010) 'Kant's First Paralogism'. *Philosophical Review* 119 (4): 449–495.
- Rödl, S. (2013) 'The Single Act of Combining'. *Philosophy and Phenomenological Research* 87 (1): 213-220.
- Rovane, C. (1998) *The Bounds of Agency: An Essay in Revisionary Metaphysics*. Princeton: Princeton University Press.
- Rovane, C. (2012) 'Does Rationality Enforce Identity?' In *Self and Self-Knowledge*. Ed. Coliva A. Oxford: Oxford University Press.
- Sassen, B. (2008) 'Kant and Mendelssohn on the Implications of the "I think"'. In *The Achilles of Rationalist Psychology*. Eds. Lennon, T. and Stainton, R. Dordrecht: Springer.
- Sartre, J. (2004) *The Transcendence of the Ego*. Trans. Brown, A. London: Routledge.

- Schachter, J. (2008) 'Leibniz's "Achilles"'. In *The Achilles of Rationalist Psychology*. Eds. Lennon, T. and Stainton, R. Dordrecht: Springer.
- Strawson, P.F. (1966) *Bounds of Sense*. London: Methuen Press.
- Tester, S. (2014) 'Kant's Metaphysics of Mind and Rational Psychology'. Dissertation submitted at Humboldt-Universität zu Berlin.
- Thiel, U. (1996) 'Between Wolff and Kant: Merian's Theory of Apperception'. *Journal of the History of Philosophy* 34: 213–32.
- Thiel, U. (1998), 'Self Consciousness and Personal Identity'. *The Cambridge History of Seventeenth-Century Philosophy Vol I*. Eds. Garber, D. and Ayers, M. Cambridge, UK: Cambridge University Press.
- Thiel, U. (2001) 'Kant's Notion of Self-Consciousness in Context.' *Kant und die Berliner Aufklärung: Akten des IX. Internationalen Kant-Kongresses Vol. 2*. Eds. Volker, G.; Horstmann, R.P.; Schumacher, R. Berlin: De Gruyter.
- Thiel, U. (2011) *The Early Modern Subject: Self-Consciousness and Personal identity from Descartes to Hume*. Oxford: Oxford University Press.
- Watkins, E. (2005) *Kant and the Metaphysics of Causality*. Cambridge, UK: Cambridge University Press.
- Wuerth, J. (2006), "Kant's Immediatism–Pre-Critique," *Journal of the History of Philosophy* 44 (4): 489–532.
- Wuerth, J. (2010a) 'The First Paralogism, its Origin, and its Evolution: Kant on How the Soul Both Is and Is Not a Substance'. In *Cultivating Personhood: Kant and Asian Philosophy*. Ed. Palmquist, S. R. Berlin: Walter de Gruyter,
- Wuerth, J. (2010b) 'The Paralogisms of Pure Reason'. In *The Cambridge Companion to the Critique of Pure Reason*. Ed. Guyer, P. Cambridge, UK: Cambridge University Press.
- Wuerth, J. (2014) *Kant on Mind, Action, & Ethics*. Oxford: Oxford University Press.
- Wunderlich, F. (2001) 'Kant's Second Paralogism in Context.' In *Between Leibniz, Newton, and Kant*. Ed. Lefeuvre, W. Dordrecht: Kluwer.
- Wunderlich, F. (2005) *Kant und die Bewusstseinstheorien des 18. Jahrhundert*. Berlin: Walter de Gruyter.