UNIVERSITÀ DEGLI STUDI DI MILANO

Doctoral Programme
Philosophy and Human Sciences
(XXXVII cycle)

DEPARTMENT OF PHILOSOPHY
"PIERO MARTINETTI"

PhD DISSERTATION

KANT ON THE INACCESSIBILITY OF NON-SENSIBLE REALITY

M-PHIL/05 — History of Philosophy

SUPERVISOR
Prof. Stefano Bacin

PhD COORDINATOR
Prof. Niccolò Guicciardini Corsi Salviati

PhD STUDENT
Rodrigo de Araujo

Academic Year 2023/2024

Acknowledgements 3

Introduction.....4

1 Individual Noumena and the Limits of Discursive Understanding.....9

1.1 Thoroughgoing Determination.....10

1.2 Criteria for Objective Reference and Numerical Distinctness.....18

1.3 Intellectual Intuition.....25

1.4 The Positive Use of the Notion of 'Noumenon' 29

1.5 Categories and Principles Beyond Experience 33

2 Things in Themselves in the Amphiboly 39

2.1 Things in Themselves as Intrinsic Natures?.....39

2.2 Kant's Overall Argumentative Strategy in the Amphiboly and its Application to the Inner/Outer Dichotomy 42

2.3 Relational Properties Without Intrinsic Natures: A Logical Impossibility? 47

2.4 Kant's Agnosticism About Things in Themselves in the Amphiboly and the Problem of Identity 52

3 What I am Aware of in Pure Apperception 63

3.1 The Principle(s) of Apperception.....64

3.2 Apperceptive Self-Awareness and the 'I'.....67

3.3 What Pure Apperception is Not.....74

4 The Specter of Monadology: Kant's Reappraisal of the Concepts of Simple and Substantiale 85

4.1 Prelude: Reason, PSR and the Unconditioned.....86

4.2 Kant's Ambiguous Statements on Monadology and Simplicity.....96

4.3 The Simple as Unconditioned 100

4.4 On the Analytic Truth of Monadology 107
4.5 The Substantiale as Unconditioned 116
4.6 Monadology, Substantiality and the Soul 127

5 Transcendental Idealism Pushed to its Limit: The Existence of Noumenal Reality 134

5.1 The Import of my Deflationist Account 134
5.2 The 'Absurdity Claim' 137
5.3 Concluding Considerations: Kant's Story about Affection 147

Abbreviations of Kant's Works 157

Bibliography 158

Acknowledgements

There are many people that I would like to thank for many different reasons, but I will limit myself to a few in order to keep it as brief as I can.

This work would not have existed without the support of professor Stefano Bacin during these years as my supervisor. More than that, though, he encouraged me to continue my research and pursue a PhD when, towards the end my master's, I was at the crossroads and was not convinced that I could actually make it.

I am grateful to professors Tobias Rosefeldt and Katharina Kraus for kindly hosting me and discussing together at Humboldt Universität zu Berlin and Johns Hopkins University respectively. My time in Berlin and Baltimore was great and especially fruitful for my research.

Among the many people with whom I had the chance to discuss about Kant, a special thanks goes to Lorenzo Spagnesi, who has also been patient enough to read more than half of my dissertation and to provide his useful remarks.

Thanks to my friends from the PhD room for making my days at the university light and fun, but also for sharing their research and for their help with mine in many occasions.

Needless to say, I am always grateful to Mom and Remo for supporting me and my choices in every possible way.

Last but not least, a special thought goes to Benni: if the last years have been the best of my life so far it is largely because of her.

Introduction

Kant's transcendental idealism, and in particular his stance towards noumena, has been matter of controversy and discussion since the publication of the *Critique of Pure Reason*, and today it is without a doubt a major topic in Kant scholarship. Along with the pressing philosophical relevance of similar issues, the exceptional quality of contemporary literature is another reason why I decided to take part in the debate.

I spare the reader an overview of the contemporary debate on Kant's transcendental idealism, since excellent ones are already available[1]. The position I aim to defend here, with some crucial caveat, can be considered overall *deflationist*. By that term I aim to characterize my view in opposition to those which are often labeled as '*metaphysical*' readings of Kant. The latter label includes of course the most various kinds of interpretations, and the only thing I take them to have in common is that they ascribe to Kant some level of insight into certain features of non-sensible reality. Throughout the chapters of the present work, I will try to argue that Kant is deeply agnostic about non-sensible beings in respects that most interpreters fail to appreciate. Most of my efforts will thus be directed at showing what, according to Kant, we cannot say about noumena, and why does he establish such restrictions.

Chapter 1 poses the basis for my deflationist account by showing that our very conception of non-sensible being is crucially defective from the start. For given the features of discursive understanding, not only we are never in the position to assess whether a concept successfully refers to one (or more) individual non-sensible being(s); more radically, we do not know whether such reference is a real possibility at all. Accordingly, we are not aware of whether there are any criteria whatsoever for distinguishing individuals if these do not belong to the spatiotemporal domain. The logically possible notion of intellectual intuition, while helpful, does not bring us much further in that regard. With that in mind, I thus propose a reading of Kant's distinction between a negative and a positive use of the notion of 'noumenon'. On my reading, a positive (i.e. unwarranted) use of that notion

[1] See for example Schulting (2011) and Schafer/Stang (2022); but also Allais (2015, Part One) and Jauernig (2021, ch. 1).

is made when one claims that a certain concept has one or more non-sensible beings as its objective referent(s)

Chapter 2 engages with the Amphiboly section of the *Critique* in order to see what kind of picture of things in themselves emerges there. I focus on the inner/outer dichotomy as I provide arguments against those interpreters who take such distinction to overlap that between things in themselves and appearances. It shall emerge from my analysis that, far from holding that all things necessarily have an intrinsic nature allegedly identifiable with their 'in-itself aspect', Kant is rather especially agnostic about things in themselves in the Amphiboly. In the last section of the chapter I argue against the possibility of making identity-claims involving things in themselves and their relation to appearances.

Chapter 3 examines the ontological status of the 'I' of which I am aware of in pure apperception. The most pressing question for my purposes is whether pure apperception amounts to the awareness of myself as noumenon, as some interpreters maintain. If the latter were correct, this would pose a serious threat to my deflationist account of Kant's stance towards noumena. On the contrary, I suggest that similar interpretations of pure apperception must be rejected and, in arguing so, I present my own reading. The core of the proposed view is that pure apperception is an act through which I am aware of myself as instantiating a 'logical self' to which I must be able to ascribe my thoughts if cognition is to obtain. The result of the analysis carried out in this chapter shall thus shed some light on that complex notion and at the same time bring further evidence in support of my overall deflationist account.

Chapter 4 assesses Kant's complex attitude towards monadology. Indeed, Kant makes a number of ambiguous claims on that regard that make it seem at times as if he endorsed some sort of 'noumenal monadology'. Hence the question that guides my investigation is whether Kant takes in fact non-sensible reality to have such specific metaphysical structure. The chapter starts with a preliminary consideration of an issue that I take to have major implications for the subsequent sections. That is, I argue that the Supreme Principle of pure reason is not constitutively valid of noumena, and consequently that one is not justified in assuming the existence of something unconditioned in the noumenal domain. I then turn to the concepts of

<simple being> and subsequently to that of <substantiale>. Against rival interpretations, I argue that Kant is not committed to the existence of simple beings and substantialia in the non-sensible domain (nor in the empirical one, as it is important to stress with regard to the substantiale). More than that, the proposition stating that noumena must ultimately be grounded in simple beings and/or substantialia is not even analytically true, unless the further unwarranted assumption is made to the effect that noumena must conform to the Supreme Principle of pure reason. Once again, in line with my deflationist view, what emerges is therefore that for Kant we have no insight into the metaphysical structure of a non-sensible world. However, we learn at the same time that the notions of simple being and substantiale have a positive role. These are the merely negative concepts of an unconditioned in the series of composition and inherence respectively. Hence, in its metaphysical quest for the unconditioned, human reason inevitably comes across these logical possible concepts whose real possibility cannot be demonstrated nor, however, disproved. Despite the inaccessibility for us of unconditioned conditions, reason is at least able to entertain logically consistent thoughts about a non-sensible level of reality underlying the sensible one, a realm (the former) where unconditioned conditions may for all we know find their place. Reason's metaphysical quest is thus not entirely unjustified. I conclude the chapter by showing that the notions of simple and substantiale converge in the conception of a monadic soul, and suggesting how the latter enables us to re-think the noumenality of the soul in deflationist terms.

Throughout chapters 1 to 4, though, a crucial question is purposefully avoided: Is Kant committed after all to the existence, if not of individual noumena, at least of non-sensible reality in general underlying appearances? This is the topic of Chapter 5, which concedes a positive answer to that question. I focus here on what I call 'Absurdity Claim', namely Kant's claim to the effect that the concept of 'appearance' implies something that appears. I argue that Kant's formulations of that claim are deeply unsatisfying, and that the metaphysical relevance of what is otherwise a mere conceptual entailment hinges on a much pressing issue, that is: Why should we call empirical objects 'appearances' in the first place? At a closer inspection, Kant does not provide an explicit and satisfying solution to the question, as one

might instead expect. After providing some reasons why Kant endorses this and other connected assumptions that lead him to accept the existence of non-sensible reality, I suggest that Kant's position on the matter is ultimately unwarranted and arguably incompatible with some of his radical teachings about the limits of human discursivity. I also believe, though, that what I call the 'logical possibility+' — i.e. logical possibility plus theoretical indemonstrability of the real impossibility — of non sensible beings is a major outcome of the *Critique* and would have been sufficient for most of Kant's purposes.

From this brief overview, it should have emerged how my reading of Kant's transcendental idealism differs from the main interpretative alternatives. First of all, I believe that Kant is not in the position to establish the precise relationship between the two levels of reality in terms of 'two-aspects' nor 'two-worlds'[2]. Secondly, not only Kant's commitment to non-sensible reality is, on my account, maximally general and cut to the bone when compared to the picture of current 'metaphysical' readings; also, unlike most 'metaphysical' interpreters, I do not think that Kant's commitment is entirely defensible by his own lights. Yet, as it will be evident in the chapters that follow, I am deeply indebted to readings belonging to that interpretative direction, inasmuch as they take into serious consideration those seemingly metaphysical aspects of Kant's philosophy often overlooked throughout the years. One may in fact note that, for being someone who proposes an overall deflationist interpretation, I spend some words on that which, I suggest, cannot be the object of truth-apt claims. Yet I take this to match a distinctive feature of Kant's philosophy: we constantly end up entertaining thoughts about a non-sensible reality, and we do so not only to draw the limits of human discursivity, but also, as already noted, in consequence of reason's unavoidable strive for metaphysics, not to mention reason's practical interest.

---

[2] On this point I am thus sympathetic with Gardner (1999, ch. 8) and Stratmann (2018). It must be said as well that the very labels of 'dual-aspect' and 'two-objects (or two-worlds) view', while useful to navigate in the ongoing debate, often prove to be imprecise. For instance, dual-aspect readers at times deny a strict numerical identity between appearances and things in themselves. On the other hand, proponents of two-object (or two-worlds) views often concede that we cannot properly speak of a noumenal *world*, or even of non-sensible *objects*. Overall, on a closer inspection, I have the impression that metaphysical dual-aspect and two-object views often turn out to be much closer than their proponents are ready to admit.

There is no problem with entertaining such thoughts, though, as long these are logically consistent and, most crucially here, we do not take them to be objectively, i.e. descriptively valid on theoretical grounds. A final distinctive element of my 'deflationist' account is that, since I eventually admit Kant's assumption of non-sensible reality, I distance myself from interpreters which believe that such an apparent assumption can be entirely explained away by deflationist means.

Let me conclude with a quick general remark on the present work. It was inevitable to leave aside many issues that are in fact of no secondary importance for a complete assessment of Kant's stance towards noumena[3]. This negligence is due to reasons of space, time, and sometimes to the fact that I do not have a clear enough position worth presenting. Also, some views I present would have certainly deserved a deeper development. I reserve this goal for the future.

[3] To offer some examples, I do not engage with recent (and outstanding) adverbial readings of the expression '*an sich*', as in '*Ding an sich*' (see Wolf 2023 and Rosefeldt 2024); I do not give the notion of '*Fürwahrhalten*' the deserved attention that it started to have in recent literature (e.g. Chignell 2007, Stang 2016, and Schafer 2023); I do not attempt at a systematic account of the ideas of pure reason; nor do I expand my considerations outside the theoretical domain.

1
Individual Noumena and the Limits of Discursive Understanding

It is relatively uncontroversial among scholars that, according to Kant, we cannot have determinate theoretical cognition of noumena. In this chapter I shall argue that Kant's account of the limits[4] of human understanding allows us to draw a more radical conclusion: the very notion of an *individual* 'intelligible being' lies beyond human comprehension. I further suggest that this claim best conveys Kant's ban of a positive use of the notion of noumenon. My investigation is guided by a question: Is it really possible for concepts to successfully refer to an individual object in absence of sensible intuition? Or, to put it differently, is it possible to single out one or multiple individual noumena (and not mere concepts thereof) falling under a certain concept? I shall argue that, while we cannot exclude this possibility, we cannot comprehend (and thus admit) it either.

My overall argument goes as follows. When assessing possible criteria for objective reference and individuation in absence of sensibility, it is worth taking into consideration a seemingly promising model, namely that of complete conceptual determination. Hence, section 1.1, discusses Kant's Principle of Thoroughgoing Determination in light of his characterization of concepts as intrinsically general. A first negative result is then established to the effect that the existence of an individual intelligible being would not follow from the hypothetical complete determination of its concept. In section 1.2, I further suggest that thoroughgoing determination also plays no

[4] A brief remark on my usage of the term 'limits'. In §57 of the *Prolegomena*, Kant submits the distinction between limits (*Schranke*) and boundaries (*Grenze*). Limits «are mere negations that affect a magnitude insofar as it does not possess absolute completeness» (4: 352); boundaries, in contrast, presuppose something lying outside of them. In his analysis of this distinction, Howard (2022) stresses that «[s]ensibility and the *empirical* understanding, for us mere humans, have limits and not boundaries»; instead, «*pure* understanding, however, like reason, has a boundary, because it has two distinct objects: it can employ the categories not only in cognition but also in thought» (Howard 2022, p. 9). Things are however blurrier than that. In the following, I will talk about 'limits' of understanding, since the possibility of setting a boundary between cognitions and mere thoughts of pure understanding is first made necessary by the intrinsic limits of human understanding in general, which, for instance, can neither itself intuit, nor does it know how to single out an actual object in absence of intuition.

constitutive role in Kant's general positive account of objective reference and numerical distinctness, which is crucially dependent upon spatial location instead. Since the model of a thoroughly determined concept cannot enrich our comprehension of what an intelligible being could be, section 1.3 considers a notion that is introduced precisely to provide such enrichment: the notion of 'intellectual intuition'. However, while the latter is a useful, hypothetical, ad hoc model of non-sensible givenness, we are still left with no criteria for making sense of numerical distinctness among noumena. In section 1.4, I show that these results account for Kant's ban of the positive use of the notion of noumenon. Finally, section 1.5 completes the picture by arguing that our comprehension of what a noumenon would concretely amount to is further impeded by Kant's views on the objective validity of the categories and by the restricted validity of the synthetic a priori principles established in the Transcendental Analytic. Overall, although the existence of a non-sensible level of reality in general cannot yet be excluded, from the analysis undertaken in this chapter it shall emerge how entertaining thoughts about non-sensible individuals is even more problematic for our discursive understanding than Kant's interpreters are usually willing to acknowledge.

1.1 Thoroughgoing Determination

In order to account for the limitations faced by discursive understanding when attempting to make sense of individual non-sensible beings, it is important to examine more closely some crucial features of such discursivity.

The discursivity of human understanding is an essential feature of Kant's epistemology. In the *Jäsche Logic*, this feature is described as follows: «From the side of the understanding, human cognition is *discursive*, i.e., it takes place through representations which take as the ground of cognition that which is common to many things, hence through *marks* as such» (*JL* 9: 58)[5]. Kant then goes on to add that «[e]very mark may be considered from two

[5] Similarly, Kant writes in the *Critique* that concepts are related to objects only mediately, «by means of a mark [*vermittelst eines Merkmals*], which can be common to several things» (A320/B377).

sides: *First*, as a representation in itself». For instance, the concept <personal computer> can be considered in its specificity, as we readily grasp what it means. A personal computer is something that has specific features, certain predicates contained *in* its concept which make it distinguishable from similar type of objects. However, a mark must be considered also from another standpoint: «*Second*, as belonging, as a partial concept, to the whole representation of a thing, and thereby as ground of cognition of this thing itself» (*JL* 9: 58). In this second sense, a concept is a genus and is thus broader than other concepts standing *under* it[6]. For example, the predicates included in <personal computer> are in turn concepts as well, which however are broader than that of <personal computer>: a personal computer is a solid, electronic, interactable physical object and so on. Precisely because of their generality, broader concepts can be used as the ground for cognition of the narrower ones. Kant's famous example is the empirical judgement 'all bodies are divisible': the concept <divisible> is broader than that of <body>, and because of that the former can be ascribed to the latter to generate empirical cognition. Yet it is significant that the very concept of <personal computer> is itself a genus that has many concepts under it, e.g. <laptop> or <desktop>, that has this or that specific operative system, brand, model and so on. The two ways of considering a mark or concept are not alternative, but rather two sides that belong to *every* concept. However, generality is the most distinctive aspect of concepts, for it is a result of the process of empirical concept formation. That process in fact involves the stages of *comparison* among various representations, *reflection* on their similarities, and *abstraction* of a common trait, i.e. a concept[7]. More broadly, any concept is a representation produced by our spontaneous faculty for thinking, i.e. the understanding, and «understanding can make no other use of these concepts than that of judging by means of them» (A68/B93). The result is that, while it is certainly possible to consider concepts as relatively enclosed representations, they are essentially «predicates of possible judgments» (A69/B94). Every concept is a

---

[6] That which, in Kant's terminology, is contained '*in*' a concept can be translated in contemporary terms as the concept's *intension*; accordingly, that which for Kant stands '*under*' a concept is what we now call its *extension*.

[7] See *JL* 9: 94-5.

general mark which is common to many things, and precisely because of that it can be used in judgements as a predicate for other concepts under it.

Now, a first problem arises for human discursivity: given that concepts are structurally general, that is, they are marks which are common to more than one thing, and given that cognition is expressed in judgements, where a subject is determined by means of a concept that serves as the predicate, how is it then possible to achieve a complete cognition of a thing in its qualitative particularity? This question is implicit in the chapter of the Transcendental Dialectic on the Ideal of Pure Reason. While the ultimate goal of the chapter is to critically assess the classical arguments for God's existence, in the first and especially in the second section can be found obscure but valuable claims concerning the nature of human discursivity. Kant begins Section Two by stating the so-called «principle of determinability» (PD), which is a merely logical principle resting on the principle of contradiction: «Every **concept**, in regard to what is not contained in it, is indeterminate, and stands under the principle of **determinability**: that of **every two** contradictorily opposed predicates only one can apply to it» (A571/B599). If I consider a pair of contradictory predicates which have not yet been included in a concept, it is possible to add only one of them to the content of the latter, which becomes thus more determined. This can be done at any moment for any pair of contradictorily opposed predicates taken into consideration, which means that a concept can always be further determined. However, a concept is always indeterminate with respect to those infinite predicates[8] that have not yet been considered with respect to it. All this is not per se problematic if the subject to be determined is in turn a concept, and thus general — and in fact PD regards «[e]very **concept**».

However, a different challenge is presented by the fact that a finite set of general marks cannot amount to the complete cognition of a particular *thing*. In order to solve this problem, Kant introduces, right after the brief mention

[8] It seems unproblematically correct for Kant that predicates are potentially infinite. In the *Vienna Logic*, for example, we read: «The series of coordinate marks can go on to infinity. E.g. Gold is heavy, extensible, refractory, does not rust, etc., etc. There is still not completeness here. For one could discover 1000 more such marks» (24: 834).

of the PD, the principle of thoroughgoing determination (PTD)[9]. This second principle is not merely logical but transcendental, and has more complex implications. It states: «among **all possible** predicates of **things**, insofar as they are compared with their opposites, one must apply to it» (A571-2/ B599-600). First of all, the PTD does not just concern the relation of the subject with *any* given pair of contradictory predicates that can be taken into consideration at a certain moment. It rather involves *all* possible predicates. Secondly, unlike in the previous principle, the subject of predication involved here are *things* and not mere concepts. Now, it is arguably complicated to follow Kant's shift from the logical talk of 'concepts' in the formulation of the PD to the ontological talk of 'things' in the PTD, while continuing to speak in both cases of predicates of judgements. Trying to make order, though, the PTD can be understood as the claim that the cognition of a thing as qualitative distinct from others entails the reference to — paraphrasing La Rocca (1999)[10] — an *x* with regard to which the attribution of a predicate or its opposite is already decided in advance. Seen from the conceptual perspective, it can also be said that, if such cognition were to obtain, the thing's qualitative determinacy would be caught by a completely determined concept.

Kant's next step is noting that the PTD, which is synthetic and not analytic, presupposes, as an a priori condition of a thing, the sum of all possible predicates. The fact that, in light of the PTD, things must already be assumed as completely determined before our judgements about it implies that the totality of predicates is also given as «a transcendental substratum, which contains as it were the entire storehouse of material from which all possible predicates of things can be taken» (A575/B603). It is clear, however, that this totality could not be accessible as such to us. Furthermore, the

---

[9] I follow here Guenova's suggestion: «On my reading, the principle of thoroughgoing determination addresses a crucial problem that emerges out of Kant's most basic views regarding the nature of empirical concepts: namely, the gap between the generality of our concepts and the particularity of the empirical manifold. To be able to regard our general empirical concepts as providing us with a genuine cognition of a particular object, we must be able to assume the possibility of constructing a complete concept» (Guenova 2013, p. 337).

[10] «[L]a condizione per pensare qualcosa come individuo *dato* è che […] l'attribuzione di un predicato o il suo opposto […] sia pensata riferita a un x rispetto alla quale è già decisa» (La Rocca 1999, pp. 190-1).

process of determination of a thing in judgement with respect to the infinite possible predicates is also impossible for us to complete. The totality of possible predicates is enclosed in Kant's notion of an «All of reality (*omnitudo realitatis*)» (A575-6/B603-4), which is at the same time the abstract thought of a completely determined concept which contains, for every pair of contradictory predicates, «that which belongs absolutely to being» (A5761B604), i.e. it contains all positive reality[11]. However, this is a mere idea of reason, or rather, more precisely an **ideal**, which Kant describes as an idea «*in individuo*» (A568/B596)[12]. Kant's attempt to converge the notion of a thoroughly determined concept and the thought of an all of reality is not easy to follow, and the same can be said about the subsequent hypostatization of the ideal in the *ens realissimum*, which in turn leads him further to the refutation of the so called ontological argument. In the present context, however, I am only concerned with the characterization of human discursivity that emerges from the text.

What has been said so far is summarized by the following passage:

> in order to cognize a thing completely one has to cognize everything possible and determine the thing through it, whether affirmatively or negatively. Thoroughgoing determination is consequently a concept that we can never exhibit *in concreto* in its totality, and thus it is grounded on an idea which has its seat solely in reason, which prescribes to the understanding the rule of its complete use (A573/B601).

The ideal is not descriptive of reality, but rather merely apt to a regulative use by reason in directing the activity of understanding, namely the activity of cognition through conceptual determination in judgement. The ideal

---

[11] Kant's claim seems to echo §807 of Baumgarten's *Metaphysica*: «All realities are truly positive, and no negation is a reality (§36). Therefore, if all realities are indeed maximally joined together in a being, no contradiction would ever arise from them (§13). Therefore, all the realities in a being are compossible. Now, the most perfect being is the most real of beings (§806). Therefore, the totality of realities <*omnitudo realitatum*> belongs to the most perfect being, and indeed the greatest realities that are possible in any being». I will gloss over the problem of 'real repugnance' between some of the predicates which, as Kant writes, belong «absolutely to being». See Willaschek (2018, pp. 227-8) for a brief discussion of the issue with respect to the Ideal, and Chignell (2010) for an overview on real repugnance in Kant.

[12] The sage, for example, is said to be an ideal of virtue (A569/B597).

serves as «*prototypon*», maintains Kant, of which all other things are to be considered as «defective copies (*ectypa*)» (A578/B606), for reason employs it regulatively both as the material condition of all objects of cognition, and as the archetype of a completely determined concept. I believe that Willaschek (2018, section 8.1.2) is right to read the PTD in light of his distinction between regulative and constitutive use of transcendental principles of reason. Employing a principle constitutively means taking it to be a true descriptive claim about reality; while a regulative use adopts it only as a hypothetical assumption[13]. Only the second use is warranted for theoretical reason. In light of the impossibility for our discursive understanding to achieve complete conceptual determination, a regulative employment of the PTD, which means assuming that things already have complete ontological determinacy, is necessary in order for the understanding to progressively increase its cognition of particular beings by means of general concepts[14]. The same can be applied to the related assumption that the totality of predicates is indeed given as well.

The problem of complete determination can be considered from a different but related perspective in Kant's discussion of the principle of *specification* in the Appendix to the Transcendental Dialectic. As we learn there, in order for reason to pursue the goal of systematizing empirical cognitions, concepts can, and indeed must be arranged in a hierarchical structure of genus and species. Together with the principles of homogeneity and continuity, the

[13] Needless to say, there is much disagreement among scholars on what the regulative use of ideas and principles of reason amounts to be (see also note 15 below). To give one example, some interpreters (e.g. Chignell 2007, Stang 2016 and Schafer 2023) would rather speak here of 'theoretical' or 'doctrinal' *belief* in the validity of those principles. I remain neutral on whether such mode of 'assent' (or '*Fürwahrhalten*', as Kant calls it in Section Three of the Canon of Pure Reason) is as crucial here and elsewhere as some believe (I will spend a few more words on this notion in section 4.1).

[14] Guenova suggests that the role of PTD is to legitimize our attempt to cognize by means of general concepts, in the sense that this principle entails that «we have the right to» «use our current general empirical concepts in place of the complete concept of the object» «because our general empirical concepts would form part of the content of the complete concept of the object» (Guenova 2013, p. 342).

principle of specification serves as guide in the development of a system[15]. This law prescribes reason to explore the manifoldness and variety of things belonging to the same genus. While our present system of knowledge arbitrarily stops somewhere on both sides of the chain of genera and species, it must always be possible at any given moment to go on and discover/ define a new species (as well as a new genus). Hence no species can ever be found that it is not in turn a genus, i.e. no concept that could not be a possible predicate in a judgement. As Kant writes:

> every **genus** requires different **species**, and these **subspecies** […], reason demands in its entire extension that no species be regarded as in itself the lowest; for since each species is always a concept that contains within itself only what is common to different things, this concept cannot be thoroughly determined, hence it cannot be related to an individual, consequently, it must at every time contain other concepts, i.e., subspecies, under itself (A655-6/B683-4).

'No species can be the lowest' means that concepts are always general and indeterminate with respect to some possible predicates. Reason must assume that no concept's content is ever completely determined. The descent into the chain of genera and species by means of an increasing specification is therefore a never-ending path to complete determination, which actually fosters the continuous enrichment of our system.

Now, in order to see how the just discussed account of human discursivity relates to our comprehension of the notion of 'noumenon', I would like to

---

[15] I will not engage here in detail with the problematic distinction introduced by Kant between a logical and a transcendental version of these principles. A groundbreaking account of the distinction between logical and transcendental principles is developed by Grier (2001). For a very recent, but interpretatively more pessimistic discussion of this problem in relation to the case here at issue, see Proops (2021, ch. 17). However, in this case as well I am sympathetic to Willaschek's account (2018, in particular ch. 4 and section 8.2.1), which provides an extremely detailed account of the differences between logical principles, the already mentioned regulative (and thus legitimate) use of transcendental principles and finally a constitutive (and therefore unwarranted) use of the latter. In sum, to the extent that it is relevant for the present context, it suffices to note that, while the logical principle of specification concerns concepts as such, the transcendental version of this principle is a maxim that assumes that further subspecies are given as corresponding objects in nature. Reason falls then into error when it makes a constitutive rather than a regulative use of this maxim.

focus on the following questions. Let us suppose that our understanding were able to form a thoroughly determined concept: (1) Would the formation of such concept already entail the actuality of an object that instantiates it? (2) Would it be possible for more than *one* object to fall under the same thoroughly determined concept? The two previous questions are especially meaningful here in light of a third one: (3) Can we meaningfully describe individual noumena, i.e. beings of the understanding, as nothing but thoroughly determined concepts? The latter question is what ultimately matters the most for us here, but before an answer can be provided, further considerations are in need.

Starting from question (1), I believe the answer is negative. We know that for Kant actuality does not follow from the concept of <*ens realissimum*>, and since the latter would be a thoroughly determined concept, it results that actuality does not necessary follow from thoroughly determined concepts. One might reply here that there is a difference between the higher order concept of a <thoroughly determined concept>, and the very concept itself that happens to be thoroughly determined. In the same way, we can distinguish between the concept of a <contradictory concept>, which is not itself contradictory, and a concept which instead *is* contradictory. One could then argue that, while actuality does not follow from our concept of a <thoroughly determined concept>, it may follow from a concept that we were actually capable of thoroughly determining. Yet this distinction does not seem to be part of Kant's strategy here. For if he had that in mind, in his attack at the 'ontological proof' Kant would have appealed to the aforementioned distinction in order to deny that the rationalist really does have the concept <*ens realissimum*> from which to derive the existence of God. Kant's famous point is rather that existence is not a predicate; that a concept remains the same regardless of whether an existing being instances it or not. Therefore, it seems safe to say that the view expressed by Kant here is that, even in the case of a thoroughly determined concept, its logical possibility (i.e. the lack of internal contradiction) is not sufficient to prove the

real possibility of a being that instances it (let alone its actuality)[16]. Furthermore, in the specific case of the *ens realissimum*, it is clear that Kant is not committed to its actuality (at least not on theoretical, objective grounds), for all reason needs is to make a regulative use of this idea: «reason does not presuppose the existence of a being conforming to the ideal, but only the idea of such a being, in order to derive from an unconditioned totality of thoroughgoing determination the conditioned totality, i.e., that of the limited» (A577-8/B605-6; see also e.g. *Refl.* 5759, 18: 346). Let us now switch to the two remaining questions.

1.2 Criteria for Objective Reference and Numerical Distinctness

We have just seen how for Kant the thought of a completely determined thing is an ideal which we must assume regulatively and merely on the ground that we must seek an increasing conceptual enrichment in the enterprise of accounting for the inexhaustible variety presented to us by the senses. We must thus assume, although merely regulatively, that reality is graspable by means of infinite conceptual determination. Furthermore, this demand is in some sense counterbalanced by the principle of specification in the Appendix, according to which we must always assume that a further species can always be discovered, i.e. that there are no lower species. These and other considerations suggest a negative answer to (1): complete determination of a concept would not entail the actuality of a corresponding object. Issue (2) is more complex: Would it be possible for *multiple* objects to fall under the *same* thoroughly determined concept? Answering this question requires us to discuss the general condition that allows us to distinguish between the multiple objects that may fall under a concept.

---

[16] Kant explains that something is logically possible if «the concept (thought) is possible» (B302-3), whereas establishing its real possibility means being able to refer such concept to a given object (cf. BXXVIn). Throughout the present work I shall make an extensive use of Kant's distinction between logical and real possibility, to which I ascribe much importance. However, nowhere in this work do I offer an overview on these elusive notions, which I assume the reader to be at least familiar with (for a detailed analysis, see Stang 2016). Some brief clarificatory remarks, though, will be made on specific occasions, e.g. in section 4.1 and 5.1.

It is useful to briefly contrast Kant's and Leibniz's views on these issues[17]. Although it is a complex matter of interpretation[18], it can be said that for Leibniz those concepts which are not abstract and which therefore do correspond to an individual substance are in principle completely determined[19]. While our finite mind is not able to achieve complete conceptual determination, we know that God instead is able to do that. Indeed, in order to create the best of all possible worlds, God examines the complete concept of things and decide which ones to actualize. Hence, despite our epistemological constraints, we know that the world is constituted by substances which find perfect correspondence in the thoroughly determined concepts that God is able to think. The role of unreachable *prototypon* played by the complete concept is here derived from our knowledge of an inner structure of reality — i.e. from the monadology — plus that of God's faculties. It seems correct to claim, as Friebe (2022) does, that for Leibniz a concept (or at least some concepts) *analytically* entails its objective referent. Things are certainly made more complicated by the fact, among others, that it is up to God to actualize certain concepts. For this implies some kind of hiatus between the concept and its actuality as a substance. Without digging too deep into Leibniz's own complex theoretical framework, though, it is fair enough to maintain that, for him, by means of conceptual analysis it should be possible to reach a *unique* individual, which is thereby qualitatively distinct from any other being. A unique individual is completely exhausted by its thoroughly determined concept. Therefore, rephrasing the principle of the identity of indiscernibles, two different objects cannot fall under the same thoroughly determined concept. In this way, numerical distinction and objective reference is established by means of pure understanding alone.

For Kant things stand quite differently. Kant's individual object is not unique. As he argues in the Amphiboly, two individuals, as appearances,

---

[17] In fact, that is what Kant himself does in the Amphiboly chapter. The argument in the Amphiboly will be discussed more in detail in the next chapter.

[18] See for example Di Bella (2013) for a distinction between two different positions, that can be found in Leibniz's texts, concerning the completeness of concepts.

[19] See, for example, Leibniz's *Discourse on Metaphysics*, §8.

may have the exact same conceptual properties. They would nonetheless be numerically distinct in case they occupy two different regions of *space* (at the same time). Location in space is then the criteria for individuation, and is therefore required if our discursive understanding is to apply a concept to a real object. For example, the concept <table> can have two or many more objective referents in space. The two objects in front of me can perfectly be two instances of the concept <table>, independently of the generality of the latter — and this of course in a completely different sense from that in which two subspecies can be said to instantiate their common genus. Therefore, while it is possible to search for a more determined concept that better expresses the particularity of a certain object, this is *not* required for us to be able to apply a concept (e.g. <table>) to an individual, and furthermore does not prevent us from picking out more than one spatiotemporally given individual exemplifying the same concept. For instance, each of these two tables in front of me is not, and need not be the unique referent of the concept <table>.

Now, while this makes it clear that thoroughgoing determination is not necessary for objective reference, the question is whether it is sufficient to single out one individual object. The question is extremely relevant, for if this were in principle possible then it would also be possible to make sense through pure understanding alone of what it means to be an individual. That is, an individual would be defined as the object of a thoroughly determined concept. A univocal answer to this interpretative issue can hardly be given once and for all for reasons I will discuss in a moment, but I contend that the most plausible answer is that thoroughgoing determination is *not* sufficient to single out one individual object; or, to put it differently, that *yes*, it is in principle possible for multiple objects to fall under the same thoroughly determined concept. In the *WRP* essay, Kant effectively writes that it «is a contradiction to think, by mere concepts of the understanding, of two things as external to one another […] ; *it is always merely one and the same thing thought twice over (numerically one)*» (20: 280, italics mine). The point is that, as far as we can tell, spatial location is the only criterion that we know for numerical (and not qualitative) distinctness[20]. Consequently, it would be

[20] This, again, is pointed out clearly by Friebe (2022).

already misleading to ask whether in thought alone a concept objectively refers to only one or rather more individuals. This applies indifferently to a general concept as well as to a hypothetical thoroughly determined concept, and there is no reason why in principle a complete concept could not have two objective referents occupying different regions of space at the same time. Now, it is easy to object that, if among the predicates of a complete concept of an individual substance we also count its relational properties, which include its location in space over time, then such concept would indeed refer univocally to one individual, and the principle of the identity of indiscernibles would prove to be true. This may in fact be exactly Leibniz's reply. Stefano Di Bella notes that, according to one of the two strands about complete concepts that can be detected in Leibniz, such kind of concept is

> thought of on the model of concrete individuals, where historical and relational predicates come to the fore, so that belonging to a world becomes a fundamental condition for the possibility of individual concepts. […] [This way of considering complete concepts] prefigures, in a sense, the Kantian idea of complete determination of empirical objects through their location within the whole of experience. This, admittedly, in a pre-critical vein […] (Di Bella 2013, p. 131).

The crucial disagreement turns then on whether Kant is right or not to claim against Leibniz that spatiotemporal location is irreducible to concepts. Since, however, engaging in a full comparison of the views of the two philosophers and establishing which one is superior is beyond my scope, I shall not discuss the matter further.

We just saw that spatial location is a sufficient criterion for numerical distinctness. I find it plausible that for Kant, although this sensible criterion for identity and multiplicity in general need not necessarily be the only possible one, it is nevertheless *the only criterion we are aware of*. As I anticipated, though, the answer here suggested to question (2) is not immune to doubt on a textual basis, since, unfortunately, the relevant texts are too ambiguous on the matter. We can start quoting the following passage from the Ideal chapter:

> the concept of an *ens realissimum* is the concept of an individual being [*der Begriff eines einzelnen Wesens*], because of all possible opposed predicates, one, namely that which belongs absolutely to being, is encountered in its determination (A576/B604).

The most immediate reading of it would have Kant claiming precisely what I am opposing here, namely that complete determination of a concept is sufficient to establish the singularity of its (possible) referent. However, the picture gets more complex and in fact confused if we consider some *Reflexionen* alongside with some passages from the *Lectures on Logic*.

In the latter, Kant shows an ambiguous attitude towards the term «*conceptus singularis*». I quote a passage from the 1770's *Blomberg Logic*, for it is interesting to compare what we find here with claims that appear in later lecture transcriptions:

> All *conceptus* are either
>
> A. *conceptus singulares* or
> B. *conceptus communes*.
>
> In the former I think only one thing, but through the latter I go further, namely, I think that which is common to many things.
> Thus the former concepts consider something individual. Thus, e.g., Rome, Bucephalus, etc. This is a *conceptus singularis.* […] The representations of immediate experience are all *conceptus singulares,* for they represent individual things. Mediate concepts of experience, however, which are abstracted from many experiences, are *conceptus communes;* so, too, are all our concepts of reason. The concept of God, to be sure, only belongs to one being, to which alone it is to be ascribed, but without settling whether only one God is possible or more, one can make out several concepts of God. But afterwards a proof must be added that the concept of God is a *conceptus singularis* (24: 257).

The end of the passage may seem to support the just mentioned immediate reading of the Ideal (A575/B604): God is a *conceptus singulars*, for it «only belongs to one being, to which alone it is to be ascribed». Less clear is what comes afterwards, for we read that it is not settled «whether only one God is possible or more» and that a proof must still be added «that the concept of

God is a *conceptus singularis*»[21]. Two things are noteworthy here about the way Kant describes a '*conceptus singulares*'. First of all, Kant mentions «Rome» and «Bucephalus», i.e. proper names, as examples of singular concepts. Yet when confronted with later lecture transcripts this example turns out to be problematic. At *DWL* 24: 756 we learn that besides the difference between species and genera there is a *differentia numerica*, which is expressed through singular concepts that, «[a]mong men we indicate them by *nomina propria*». However, just a couple of pages earlier we explicitly read that «Socrates is not a *conceptus*» (24: 754). Secondly, in the quoted passage from the *Blomberg Logic* Kant pairs singular concepts with «representations of immediate experience». The reader will correctly notice that such description fits intuitions better than concepts. Indeed, in later transcriptions from the critical period we read that intuition is a «*singular* [*einzelne*] representation (*repraesentatio singularis*)» (*JL* 9: 91; for similar claims, see also *VL* 24: 904, *DWL* 24: 754, 24: 756). More than that, Kant explicitly characterizes the opposition between intuition and concepts as one between singular and universal representations (e.g. *VL* 24: 805), and he goes as far as to claim that «one must not say *conceptus communis,* because this would be a tautology, because every *conceptus* is also *communis*» (*VL* 24: 905; cf. *JL* 9: 91); or that «if a representation is not a *repraesentatio communis,* then it is not a concept at all» (*VL* 24: 908). Accordingly, Kant notes that the very distinction among singular and universal concepts is misleading, since what can be in fact singular is the *use* of a concept in judgement, for instance when I apply a concept (e.g. <man>) to an individual (see *VL* 24: 908-9 and *JL* 9: 91). Despite some ambiguities, then, it emerges that for Kant a singular concept is a *contradictio in terminis*, for a proper concept can in principle be applied to multiple individuals.

It remains unclear to me what is Kant's stance in the case of the *ens realissimum*. Its peculiarity is not only that it is a case of <thoroughly determined concept>, but that it is supposedly determined always through the positive predicate among each of the infinite pair of opposites, i.e. «that

[21] God or the notion of *ens realissimum* is described as a conceptus singularis in other occasions, e.g. *Refl. 4253* (17: 483), *5759* (18: 346), and *5776* (18: 351-2). Is worth mentioning, however, that in the latter note Kant is explicitly commenting a passage from Baumgarten's *Metaphysica* (§114).

which belongs absolutely to being» (A575/B604). In a note dated between 1769 and 1771, Kant seems to suggest that this is exactly the reason why this notion must be regarded as a singular concept: «the general concept of a being that has limits comprehends many things under itself, but that of the unlimited is a *conceptus singularis*» (*Refl. 4253*, 17: 483). It must be reiterated that this note belongs to the pre-critical period; and while God or the notion of *ens realissimum* is described as a conceptus singularis in other places as well (e.g. *Refl. 5759,* 18: 346 and *5776*, 18: 351-2), these texts often engage explicitly with specific passages of Baumgarten's *Metaphysica*, or more generally with 'rationalist' arguments. It could be argued that the passage form the Ideal section belongs to the last category as well, because it reconstructs the path that leads to the arguments for God's existence. Further considerations could be made in the attempt to make sense of Kant's usage of the self-contradictory expression 'singular concept' and of its connection to the concept of *ens realissimum*. Without bringing the discussion too far from our focus here, we can conclude the following with regard to issue (2). Let us concede that the concept <ens realissimum> is supposed to pick out one single entity, that only one existing thing could fall under it. Yet the concept of such unlimited thing is, precisely as such, an exception in Kant's account of concepts. There is no reason to believe that, on Kant's account, even if our understanding were able to produce thoroughly determined concepts, these could only be instanced by one single existing being — unless the concept were determined just through (infinite) positive predicates as in the case of the *ens realissimum*. The most plausible answer to question (2) raised above therefore remains the following: there is no reason to exclude that more than one existent object could actually fall under the same thoroughly determined concept.

This is relevant because we are thus able to answer question (3), namely whether it is possible to describe individual noumena in terms of thoroughly determined concepts. I suggest that the correct answer is: *As far as we are able to tell: No*. Even when hypothetically stretching the model of discursive understanding in order to obtain the idea of an infinite (but actual) conceptual determination, we are not able to conceive alternative criteria for objective reference and numerical distinctness apart from those involving space and time — that is, besides those valid for phenomena but not for

'intelligible beings'[22]. How then can we make sense of the idea of an 'intelligible being', if these are not to remain empty words?

1.3 Intellectual Intuition

I have argued that the idea of thoroughgoing determination cannot help us making sense of the notion of intelligible being. Indeed, even if complete determination of a concept were achievable, this would not thereby entail the existence of some corresponding object. The gap, as it were, between mere thought and objective reality would still persist. Moreover, a complete determination of a concept is not itself necessary nor sufficient for us to successfully refer a concept to an individual thing belonging to the sensible domain. Kant's account of concepts is thus probably one reason why he insists in characterizing the noumenon taken in the positive sense as the object of an *intellectual intuition*. It is easy to see how the possibility of an intellectual intuition is designed to overcome the limitation of discursive understanding and make some sense of the notion of a 'being of pure understanding'. A hypothetical intuitive understanding would in fact have *immediate* access to its object through a spontaneous intellectual act, which amounts to say that the gap between pure thought and objective reality is bridged *ex hypothesi*.

However, Kant firmly denies the possibility for us to perform such an act. More than that, as Kant states, «the possibility of [intellectual intuition] we cannot understand» (B307). He in fact opposes the very conception of an intuitive intellect — often equated to the divine one — to our discursive understanding. Now, Kant mentions the idea of a non-discursive understanding in different contexts. In order to avoid further complications,

---

[22] In the Amphiboly we find a passage that may be taken to suggest that Kant accepts Leibniz's principle of the identity of indiscernibles as valid for noumena: «If an object is presented to us several times, but always with the same inner determinations (*qualitas et quantitas*), then it is always exactly the same if it counts as an object of pure understanding, not many but only one thing (*numerica identitas*)» (A264/B320). I will later (2.4) argue that such interpretation — defended, for example, by Wood (2005) — does not fit the overall argument of the Amphiboly. Here it suffices to suggest that this passage is best read as claiming that through pure understanding alone we lack criteria for distinguishing between different individuals falling under the same concept.

though, I shall leave aside here the role it plays in other contexts, e.g. in the possibility for us to formulate teleological judgements[23]. I will limit myself to a few notes concerning intellectual intuition as that faculty to which objects are immediately given in thought. A first point made by Kant when introducing this idea in the PN chapter is that we must be able to conceive the possibility in general that there could be some type of intuition that differs from the sensible one. He then identifies this alternative with the intellectual intuition[24]. I find this equation striking and not much justified. For why couldn't we mention the possibility of a non-spatiotemporal sensibility, or even the case of another kind of unknown faculty for intuition which is neither sensible nor intellectual[25]? These hypotheses are undeniably vague and barely graspable, and yet not contradictory. That said, Kant may have found it easier to focus on the traditional paradigm of a divine intellect, i.e. one in which there is no difference between possibility and actuality, for the object is produced by the very act of thinking it.

Now, Förster (2002) finds it necessary to stress the difference between intellectual intuition taken as an act which produces its object on the one hand, and as a non-sensible intuition of things in themselves on the other[26]. However, I believe that the distinction is ultimately superfluous, since the object given to an intuitive intellect is for Kant to be characterized as being something 'in-itself' precisely insofar as it is the object of a productive intellect, that is, inasmuch as it is produced in the same act through which is given. In fact, in case the object be completely produced through a

---

[23] For a broader discussion, see for example O'Farrell (1979), Tilliette (1995), Förster (2002). In particular, Förster has subtly distinguished the variety of capacities usually taken to converge in the multifarious notion of intellectual intuition. He has also noted how they did not have the same legacy in post-Kantian thinkers. While many of his considerations are indubitably valuable, I nonetheless agree with Quarfood's (2011) point that the unsettling oscillation in Kant's characterization of the intellectual intuition (and intuitive understanding) is itself indicative of how general is the notion of non-discursive understanding in Kant, which comes into play first and foremost for contrastive purposes only.

[24] «But if we understand by that an *object of a non-sensible intuition*, then we assume a special kind of intuition, namely intellectual intuition» (B307).

[25] Kant nevertheless admits *sensible* types of intuition that are different form ours (see B148).

[26] See also Haag (2015), who draws heavily on Förster's account.

spontaneous act of the understanding and simultaneously also given to it, there would be no further sense in which the object could have a constitution 'in-itself' apart from the way it is given in (intellectual) intuition. Furthermore, a sharp distinction between a kind of intellectual intuition that produces its object and one that intuits things in themselves may misleadingly suggest that the latter are entities which are actually out there only to be intuited by some non-discursive understanding. But this is not Kant's point. His point is rather that discursive understanding could never have access to an actual object without the contribution of sensible intuition — not even if it were able to actualize the infinite process of determination and achieve a thoroughly determined concept. And the only non-contradictory alternative that for Kant we can consistently think of is that of an intuitive intellect which could somehow be acquainted with and cognize a being of pure thought, and consequently a thing in itself. As a result, for all we know, a thing in itself can be an object for a subject only as the product of a hypothetical act of intellectual intuition.

However, my main point here is that, whereas intellectual intuition is conceived as a model for non-sensible giveness, a criterion to establish numerical distinction among noumena is nonetheless still lacking. For how could we establish sameness or distinctness of the content of two acts of intellectual intuition? Stang (2022) makes two related points on that regard[27]. First, he speculates that the identity among intelligible objects could be traced back to the identity of cognitive acts of intellectual intuition — although the problem, as he notes, would then be pushed back to that of finding a criterion of identity among such acts (cf. p. 313). That said, this model in general is admittedly no more than speculation, and yet, insists Stang, his aim was «merely to point out that this is a topic about which we can entertain logically consistent thoughts and examine their logical relations, even though none of our thoughts rise to the level of cognition, much less of knowledge» (p. 314). Stang's remarks do not represent a challenge for what just said, since we are left where we started: we are not aware of any criterion individuation outside the sensible domain; more than

[27] In the following, I will not argue against Stang, but merely show that the remarks he makes are not necessarily incompatible with mine, and thus do not threaten the claim I am defending here.

that, on theoretical grounds alone, we cannot even tell if any such criterion exists or not[28]. This brings us to a second point stressed by Stang. On moral grounds, we must admit that there are criteria for numerical distinctness in the noumenal realm. Because our moral belief in the highest good entails that God must «be able to cognize identity and difference *among* noumenal wills» (p. 314). While correct, this is not more scandalous — or perhaps, depending on one's view: it is just as scandalous as — Kant's claim that our awareness of the moral law compels us in complex ways to believe in the actuality of specific supersensible entities[29]. Accounting and evaluating Kant's move is a complex issue that cannot be addressed here. Nevertheless, I am convinced that the enterprise of discussing Kant's notion of noumenon without accepting his attempt to solve metaphysical problems on practical grounds, or simply the attempt of starting from what could be called his 'theoretical metaphysics', is interpretatively useful and worth undertaking from a philosophical standpoint. From such perspective then, it can be concluded that we are left with very little when trying to achieve some understanding of what does it mean for something to be a singular 'intelligible being': it makes no sense in relation to our discursive understanding and, while not contradictory, it is nonetheless barely comprehensible when problematically characterized as the object of an intellectual intuition.

---

[28] Stang also shows that Kant himself, in particular in his lectures on rational theology, suggests the following model of «[d]ivine intuitive cognition of objects»: it «appears to be a case of what is now known as "maker's knowledge": a non-passive or non-sensible cognition of an object as being created by that very condition itself» (Stang 2022, p. 325). Stang argues as well that something very similar is at work for us in the consciousness of my freedom, which «is the ground of the moral law biding me» (Stang 2022, p. 329). Regardless of this last issue, which I will not engage here, it must be noted that the model of 'maker's knowledge', even if it were correct, would not by itself provide a *criterion* for identity and sameness among noumena. For it means only that, by producing the object, the subject *would just happen to know immediately* whether this specific object is the same thing as the product of another cognitive act or not. Therefore, this model is not particularly useful for us to enrich our comprehension of the general notion of 'intelligible being'.

[29] Yet it seems correct to say that even practical considerations cannot entirely overcome our inability to make sense of what it would mean to refer a concept to a non-sensible object. Indeed, Kant writes of the three ideas of freedom, immortality and God that «we are instructed by it [i.e. «an apodictic practical law»] **that they have objects**, although *we are not able to show how their concept refers to an object*, and this is not yet cognition **of these objects**» (*CPrR* 5: 135, italics mine).

1.4 The Positive Use of the Notion of 'Noumenon'

We are now in a better position to discuss Kant's distinction between a positive and a negative use of the concept of noumenon. The concept of 'noumenon', of a non-sensible being, is empty in the Kantian sense that no sensible intuition — which is the only possible intuition for beings like us — corresponds to it. More precisely, in his table of «Nothing» at the end of the Transcendental Analytic (A290-2/B346-9) Kant reckons «*noumena*» as a case of «*ens rationis*», i.e. an «[e]mpty concept without object», «which cannot be counted among the possibilities although they must not on that ground be asserted to be impossible»[30]. Now, as long as we consider the thought of something «**insofar as it is not an object of our sensible intuition**», that is, precisely as an empty concept for which no intuition can be presented, then we have «a noumenon in the negative sense» (B307)[31]. Kant distinguishes this use from a positive one. In the same passage, indeed, he continues:

> if we understand by that an **object of a non-sensible intuition**, then we assume a special kind of intuition, namely intellectual intuition, which, however, is not our own, and the possibility of which we cannot understand, and this would be the noumenon in a **positive** sense (B307).

Yet accounting for this distinction is not so straightforward, since the passages that Kant presents us are perplexing in some respects. For instance, in the Remark to the Amphiboly, Kant mentions «objects of a non-sensible

---

[30] We are at least able to establish the logical possibility of a certain concept of *ens rationis* (e.g. of the concept of God): the logically possible concept of an *ens rationis* is therefore a *logical something*, and thus *not* nothing at all. Nevertheless, we are not in a position to prove its real possibility. In sum, the concept of an *ens rationis* is the concept of a logically possible something whose real possibility, however, we are not able to establish due to its non-sensible nature — although it «must not on that ground be asserted to be impossible». For an insightful analysis of the table of nothing in light of Kant's theory of modality, see Stang (2016, section 6.6).

[31] On the negative use of the notion of noumenon Kant also famously writes: «The concept of a noumenon is therefore merely a **boundary concept**, in order to limit the pretension of sensibility, and therefore only of negative use. But it is nevertheless not invented arbitrarily, but is rather connected with the limitation of sensibility, yet without being able to posit anything positive outside of the domain of the latter» (A255/B310-1).

intuition» by calling them «noumena in this merely negative sense», which «must of course be allowed» (A286/B342). The 'object of a non-sensible intuition' is thus designated first as noumenon in the «**positive** sense» (B307), but later as noumenon in the «negative sense» (A286/B342). Therefore, if we read these passages, as natural, as employing the notion of 'object of a non-sensible intuition' to define what a 'noumenon taken in the positive' (B307) or 'negative sense' (A286/B342) is, then we must conclude that they represent a plain textual contradiction. More in general, though, the main problem of Kant's account is the lack of clarity on whether a *positive* use of the notion of noumenon is made already by simply considering the logical possibility of an object of intellectual intuition, as B307 would appear to suggest, or rather only in case we hold that hypothesis as a real possibility. I submit that the core of Kant's distinction is best read in the latter way. When Kant emphatically claims at B307 that «we cannot understand» the «possibility» of an intellectual intuition, he must have had in mind here the latter's *real* «possibility». For if the former option proved to be correct, then, given that Kant often seems to allow for us to at least entertain the logical possibility of intellectual intuition, a positive use would turn out to be admissible after all.

Hence my view is that Kant is not presenting us with two types of entities[32]. As Allais correctly puts it, «the distinction between a positive and a negative conception of noumena is not an attempt to sort noumena into kinds» (Allais, 2015, p. 61, fn. 3), but rather, as I see it, to define two ways of using the same concept, of which only one is warranted, i.e. the negative

[32] Among others, Willaschek (2001), Messina (2020), Jauernig (2021) and Stang (2021), for example, submit what I think are defensible readings of that distinction as one among different types of objects, namely positive noumena and negative noumena. Some authors accordingly read Kant as committed to the existence of the latter, while at the same time remaining at least agnostic towards the former. However, I find it crucial that there is no explicit and decisive textual evidence that justifies taking Kant to be committed to such distinction as one about two varieties of noumena. In general, it can hardly be said that this terminology is consolidated in Kant's works. Quite the opposite, it appears in just a handful of passages which, on top of that, present the inconsistencies discussed above. In light of that, I take my proposal to be preferable because it is at least in a better position to account for the exact import of Kant's ban of a positive use of the notion of 'noumenon'.

use[33]. The best way to make sense of Kant's ban of a positive use of 'noumenon' is therefore by equating it to the following claim:

> **Kant's ban of a positive use of <noumenon>:** A positive and thus unwarranted use of the notion of noumenon is made only when, on theoretical objective grounds, I claim to have established (or be able to establish) that a certain concept has one or more 'intelligible beings' as its objective referent(s).

A brief, perhaps trivial clarification is in order here. If we claim to *know* (on theoretical grounds) that a noumenon, i.e. a certain non-sensible, individual object falling under a certain concept (e.g. <soul>) is the object of an actual act of intellectual intuition (e.g. God's), we would thus be making a positive, i.e. unwarranted use of the notion of noumenon. My point is that one should not object here that the supposed intuitive intellect (e.g. God) would not employ concepts in his cognitive act, and thus would not describe what is at work in terms of 'an individual falling under a concept', or 'a concept having an actual referent' — an expression that I have included in my description of what making a positive use of <noumenon> amounts to. For *we* would be the ones making a positive use of <noumenon>, inasmuch as *we* would be claiming that a concept of *ours* has an existing objective referent, namely, in the previous example, a noumenal soul supposedly intuited by God.

We can thus count at least these two ways of making a positive use of <noumenon>. The first is by purporting our discursive understanding to have access to intelligible beings. This contradicts the most basic tenets of the *Critique*. The second is by claiming that some intelligible being is an actual or really possible object of some act of intellectual intuition. Since we cannot prove the real possibility (let alone actuality) of intellectual intuition, such use of the notion of noumenon is unwarranted as well.

A similar reasoning is introduced by Jauernig (2021), albeit against a crucially different background. First of all, in ch. 6 of her book she reconfigures the often employed terminology of 'negative and positive

---

[33] That notwithstanding, I will argue in the next chapter that Allais' account of noumena is ultimately mistaken as she somehow makes a positive use of the notion of noumenon.

noumenon' by construing these in terms of «Kantian things in themselves» and «Leibniz-Wolffian things in themselves» respectively. On her view, of course, only the latter are target of Kant's repeated attacks. Now, in response to some exegetical difficulties, but admittedly in contrast to B307, she additionally argues that a positive noumenon (i.e. a Leibniz-Wolffian thing in itself) should be first and foremost defined as «an object that can be cognized by a pure understanding alone» (Jauernig 2021, p. 340 fn. 70). She then suggests that Kant has in fact two notions of pure understanding in mind, and, accordingly, two conceptions of a positive noumenon:

> First, there is the conception of a noumenon as an object that can be cognized by our pure understanding alone. Second, there is the conception of a noumenon as an object that can be cognized by an intuitive intellect alone, that is, as an object of a non-sensible or intellectual intuition. The negative conception of a noumenon is the conception of an object that is not an object of our sensible intuition. Based on the results of the Transcendental Analytic, we know that noumena in the first positive sense are impossible. By contrast, we have no reason to declare noumena in the second positive sense as impossible (Jauernig 2021, p. 341 fn. 71).

In sum, Jauernig identifies three kind of objects, two of which are not (namely the object of pure understanding alone) or may not be (namely the object of an intuitive intellect) really possible. By contrast, I believe that the picture I have presented is overall smoother and more representative of the struggles that, on Kant's view, we must deal with when entertaining thoughts about the supersensible. I suggested that there is just one core notion of noumenon, but alternative ways of making a positive, i.e. unwarranted use of it. To be sure, construing a noumenon as the object of intellectual intuition is not by itself unwarranted if the latter is taken to be no more than a logical possibility. In this case we would be sticking to a merely negative use of the notion of noumenon. Indeed, as I have argued, considering that we are aware of no criteria for numeric distinction within a non-sensible realm, Kant takes that construal to be useful as it is our best, and yet still deeply insufficient attempt to make some sense of something like an individual non-sensible being.

1.5 Categories and Principles Beyond Experience

At this point, one may wonder whether Kant's ban of the positive use of the notion of noumenon entails that our very ability to *think* noumena is compromised. In order to shed some light on this question, consider the complex and much debated issue of the employment of 'pure' or 'unschematized' categories in Kant. Given that categories are pure concepts, meaning that they do not by themselves contain sensible elements, it is natural to wonder what we can make of them in abstraction from sensible schemata[34]. While, to my knowledge, Kant never uses the word 'unschematized', he mentions — more or less problematically — a use of the categories in abstraction from sensible intuition in numerous places (e.g. B128-9 and A350). In the Paralogisms (see A402-3), for example, Kant explicitly mentions a «merely transcendental use of the category» and the employment of the concept of substance as a «pure intellectual concept»[35]. There he concedes that pure categories could very well be employed when I think myself as the subject of my thoughts — although he warns at the same time that in that case they would be of no use for the aims of the 'rational psychologists'. Now, interpreters have often insisted that, although we cannot rely on pure categories to cognize noumena, we are nonetheless entitled to *think* noumena by means of them. I ask here: Is that correct? In the aforementioned passages on 'pure categories' noumena are not explicitly involved. For instance, it is matter of debate among scholars whether the self which for Kant can be thought under the pure concept of substance should be characterized as 'noumenal' at all[36]. Yet there are at least two more

---

[34] For an example of the debate on the applicability of unschematized categories to noumena, see Kohl (2015) and Marshall's (2018) response to it.

[35] Focusing on the category of substance, McLear (2020a) draws a distinction between a «logical», a «metaphysical» and a «schematized» conception of substance. While his account is appealing, he refrains here from establishing which (if any) of these conceptions amounts to the 'unschematized' category of substance, and whether any of them may be deployed in producing meaningful thoughts about noumena.

[36] See Chapter 3 below.

contexts where Kant seems to allow for some kind of employment of categories to think noumena despite the absence of sensible schemata. This seems to happen in the practical domain (most famously when conceiving our freedom by means of the category of causality, see e.g. *CPrR* 5: 55-6), and in the even more problematic case of noumenal affection as the ground of appearances[37]. I will limit myself here to a couple of brief and generic remarks.

Whether we can legitimately claim to think noumena through pure categories surely depends on what we mean here by 'thinking something', and in light of what we have seen in the previous sections we can distinguish two options. First of all, it is correct to maintain that categories are necessarily involved in the act of thinking, in being intellectually aware of something as an object for us. Categories represent the conceptual framework for any thought of an object whatsoever: «I can think nothing without a category» (*CPrR* 5: 103). By means of pure categories I may thus *purport* to think noumena, i.e. I purport to represent something when attempting to provide some ontological consistency, as it were, to the logically possible concept <noumenon> (or to the concept of some specific non-sensible being, e.g., <God>)[38]. As Quarfood nicely puts it: «When we try to attain knowledge of things by mere concepts, what we obtain is, one might say, the mirror image of the structure of our discursivity» (Quarfood 2004, p. 58)[39]. Yet if, by contrast, 'thinking something' entails the additional requirement that we establish that our concepts *successfully* refer to a real object, or at least that it is really possible for them to do so, then it does *not* seem the case that we can actually 'think' noumena in this sense. For not only we cannot single out an actual object falling under this concept[40]; more than

---

[37] On the problem of affection, see Chapter 5 below.

[38] This achievement, though, is far from irrelevant for Kant. Entertaining a logically possible discourse that is supposed to be about noumena proves useful from the practical, but also from the theoretical standpoint (e.g. in solving the antinomies). More on the importance of securing the logical possibility of noumena will be said in Chapter 5.

[39] In the lectures on metaphysics we read: «We can make concepts of things in general only through the understanding, even if no object is given, because we are representing to ourselves only the manner in which we can think an object» (*MM* 29: 798).

[40] To reiterate, at least not on purely theoretical grounds.

that, what I have tried to show is that it is inscrutable for us what it would even mean for one or multiple individuals to be an instance of the concept <noumenon> (or of any specific concept of a noumenon)[41]. As we noted above, the concept of <noumenon> is 'empty', and so are unschematized categories[42]. Therefore, we should refrain from drawing strong metaphysical

[41] According to Kant, an intuitive intellect would not cognize through categories (see e.g. B145 and B148). Marshall (2018) nonetheless argues that this fact does not by itself prove that categories cannot be descriptive of noumena, for an intuitive understanding may be able to represent, if only indirectly, noumenal facts involving categorical properties. Regardless of its defensibility, what is relevant is that Marshall's argument would not affect my claim. I am not denying that categories *may* be somehow descriptive of noumenal reality; I rather maintain that we cannot establish on theoretical grounds whether this is truly the case, and neither can we comprehend what it would mean for a category to refer successfully to a noumenon. All in all, I am inclined to the opinion that the whole problem of whether categories would be meaningful for an intuitive intellect lies on the edge of non-sense.

[42] Edwards (2023) provides an enlightening and detailed account of the core difference between schematized and unschematized categories. I just recall here that the latter — as it is the case for the concept of noumenon — must be characterized as *empty*. Now, while Kant often pairs emptiness and lack of significance, it seems nonetheless clear that even empty thoughts are in some unspecified sense meaningful for him. Nevertheless, as Edwards effectively puts it, when entertaining empty thoughts, crucially, «we are not in a position to assess the thought for truth or falsity» (Edwards 2023, p. 10; for a very brief overview on the ambiguous relation between emptiness and significance, see also fn. 8 of his paper). See Roche 2010 for a more detailed discussion on sense and significance in Kant. With regard to Roche, I do not wish to take a stand on the important question investigated by him, namely what is the exact import of what he calls 'principle of sense'. Nevertheless I would like to take issue with a couple of claims by Roche that closely relate to what I am arguing in this chapter. He writes that «surely we can refer to a thing in itself. Typically, it is only a matter of abstracting from the sensible properties of the same thing as it appears, and there is presumably no special difficulty in referring to the latter» (Roche 2010, p. 677). If what he means here by reference implies is to single out an individual, then this picture is rather deeply problematic, in light of what we have said thus far (see also section 2.4 on the difficulty of talking about 'sameness' in such terms). Equally debatable is Roche's suggestion that «to the extent that [the soul] is a genuine idea (the object of which we cannot experience)» it has a referent (p. 678). For what we can truly claim is that the rational psychologist merely *purports* to refer the concept of soul to a being, and consequently, he is (and we are) 'thinking something' only in a loose sense. Instead, when a stronger sense of reference — that is most relevant for Kant — is at issue, I find that for Kant we have no theoretical tools to claim that the concept of soul successfully refers to a non-sensible entity, nor to understand how such reference would work. Logical possibility does not imply reference in this sense.

conclusions from Kant's claims on the possibility of 'thinking' beyond experience by means of pure categories[43].

I take the opportunity here to make a conclusive remark on a related, decisive issue. While we use pure categories when entertaining a logically possible discourse that is supposed to refer to noumena, the synthetic a priori principles that Kant establishes in the Analytic of Principles are crucially tied to the realm of possible experience. The Analytic of Principles is in fact one the most important outcomes of Kant's efforts in the first *Critique*, because there he takes to have exposed the synthetic a priori principles that lay the foundation for all our empirical knowledge of reality. As is well known, however, this achievement comes with a price. Those synthetic a priori principles are objectively valid as constitutive conditions of experience *in virtue* of their connection to the forms of our intuition, namely space and — especially — time. Indeed, Kant's own formulation of these principles contain some reference to sensibility. Now, categories, taken by themselves, are just concepts and do not entail any law of experience. Only the principles can inform us, for example, that changes must occur according to a causal connection, or that *realitas* has an intensive magnitude. Causal relations, for instance, are *constitutive* of all possible phenomena, and consequently of experience as a whole; but also *only* of phenomena, and thus within experience. Hence, even though we can purport to think a noumenon as a 'cause', we are not entitled to assume that a certain noumenon must be the cause and/or the effect of some other noumena[44]. Even the fact that the

---

[43] For such type of claims, which I take to be unwarranted, see for example Langton (1998), Wuerth (2014) and Allais (2015). Other authors committed as well to some kind of metaphysical reading of Kant's idealism are nonetheless more subtle on that respect (e.g. Chignell 2007 and Schafer 2022 and 2023). In his recent article on unschematized categories, Messina (2020) has written that «it would be a mistake to conclude […] that the unschematized categories are not also fundamental concepts of an object in general and so, in that sense, ontological concepts. They must be insofar as we can use them to think of noumena» (pp. 4-5). While Messina is not oblivious of the impossibility for us to establish the real possibility of those state of affairs that we would think by means of unschematized categories, claims like the one just quoted remain nevertheless ambiguous if one does not make clear whether 'thinking' must be taken to imply *successful* reference or not.

[44] As it has been argued (see Longuenesse 2015, ch. 7 and Lu-Adler 2021), this means for Kant that the Principle of Sufficient Reason is only *constitutively* valid of experience, as it becomes precisely the Second Analogy. I will later go back on the topic when discussing the supreme principle of reason in section 4.1.

moral law requires us to firmly assume transcendental freedom, thought in accordance with the pure category of cause, does not thereby commit us to a law of necessary connection of cause and effect among noumena. Quite the opposite: the logical possibility of freedom relies on that of a world where the casual law does not apply (plus the theoretical undecidability of the reality of such world). It is not my aim here to reconstruct Kant's complex arguments in the Principles chapter, but simply to stress how it makes little sense to presuppose any of the principles when entertaining logically possible thoughts about things in themselves. On theoretical grounds, no constitutive metaphysical condition can therefore be said to apply to the purported noumenal reality, nor to the relationship between the latter and the realm of experience[45]. This, I believe, is the true meaning of Kant's famous claim that «the proud name of an ontology, which presumes to offer synthetic a priori cognitions of things in general in a systematic doctrine (e.g., the principle of causality), must give way to the modest one of a mere analytic of the pure understanding» (A247/B303).

***

Summing up, what I tried to illustrate thus far is the following. It is of utmost importance to appreciate that, on Kantian grounds, our theoretical tools to even make some sense of the notion 'noumenon' or 'being of understanding'[46] are cut to the bone. For a noumenon could not be characterized as a mere concept, and, as far as we know not even a lowest species. The only way to make some sense of it is as the object of intellectual intuition, namely a faculty we do no have, that we do not comprehend, and whose real possibility (let alone actuality) we cannot establish. Accordingly, we have no theoretical grounds for claiming that a certain concept

---

[45] Although, as I concede in Chapter 5, Kant is likely inconsistent on that regard.

[46] My aim here was to argue against a certain 'hypostatized' — as it were — way of conceiving of intelligible beings. There are of course other possible ways to intend the latter notion. For instance, Kantian 'ideas', as concepts of reason which play a role in experience without ever being properly instanced in it, seem to me to offer such an alternative in some respect. One of Kant's main points in the Dialectic, as I take it, is that there is a natural tendency to hypostatize some ideas, namely transcendental ones, in our search for the unconditioned; but at the same time, also that we lack at least the theoretical tools for such a move.

successfully refers to one or multiple noumena. Finally, the synthetic a priori principles which are available to us and are constitutive of reality are rooted in the forms of our sensibility and are thus valid only within experience. As it will emerge more clearly in the next chapter, when we depart from sensibility and the criteria for individuation it provides, all we are left with are empty concepts and the principles establishing their logical relations. This is how serious we should take Kant's claim that we can make no positive use of the notion of noumenon in the theoretical domain[47].

[47] Cf. the following passage from the Remark to the Amphiboly: «we cannot thereby positively expand the field of the objects of our thinking beyond the conditions of our sensibility, and assume beyond appearances objects of pure thinking, i.e., noumena, since those do not have any positive significance that can be given» (A287/B343). Indeed, the whole page — from which I will quote again later — is particularly effective in conveying Kant's suspicion towards intelligible beings and pure categories. This is not to say, unfortunately, that every passage on similar issues is equally straightforward.

2
Things in Themselves in the Amphiboly

Right after the chapter «On the ground of the distinction of all objects in general into *phenomena* and *noumena*» (henceforth: PN chapter), Kant turns to the so called Amphiboly of the Concept of Reflection. In that section, Kant carries out a discussion which has significant consequences for our grasp of the relationship between appearances and things in themselves. In what follows, I shall approach the text by engaging with some interpreters, in particular Langton and Allais, who read the claims therein in a way that contrasts what I have been attempted to argue so far. In particular, the inner/outer dichotomy analyzed by Kant plays a significant role in their accounts. If their reading of the Amphiboly were to prove correct, that would pose a serious threat to my own account. However, I will not only attempt to show that their reading is unwarranted, but furthermore, that what Kant teaches us here goes actually in an opposite, more agnostic direction. I will start with a summary of Langton's and Allais' use of the Amphiboly for the interpretation of Kant's idealism (section 2.1). Only then will I discuss how Kant applies the overall argumentative strategy of the Amphiboly to the 'inner/outer' dichotomy and whether it is consistent or not with Langton's and Allais' interpretations (sections 2.2 and 2.3). I will conclude with some critical remarks on the alleged identity of appearances and things in themselves (section 2.4)[48].

2.1 Things in Themselves as Intrinsic Natures?

In her well-known book *Kantian Humility* (1998), Langton claims that the main thesis of Kant's theoretical philosophy, as far as concerns the relationship between appearances and things in themselves, is what she calls «Humility». Langton explains the relationship between appearances and things in themselves through the distinction between *intrinsic* and *relational*

[48] This chapter presents an adapted and improved version of Zanette de Araujo (2023).

properties of things (more specifically: of substances). This interpretation can be listed among those which in the current debate are referred to as 'metaphysical (or ontological) dual-aspect' readings. The humility at issue here thus means the constitutive impossibility for the subject to cognize the intrinsic nature of substances. Instead, as spatiotemporal appearances, things have nothing but relational properties that, on the other hand, we are able to cognize. The main passages quoted by Langton as textual evidence come indeed from the Amphiboly section. According to Langton, we can find here compelling evidence that for Kant things in themselves are substances that have intrinsic properties, and that we can cognize only the relational properties of such substances, i.e. their appearance.

Langton's book has been the target of serious criticism (e.g. Bird 2000 and 2006; Quarfood 2004, and Onof 2011). One of the reasons for such criticism is that Langton explicitly understates the role of the ideality of space and time in Kant's arguments. In doing so, she also assigns a pivotal role in Kant's system to a section that is usually taken to be rather secondary in the *Critique* — especially if we consider that in the Amphiboly Kant's main concern is to refute a few (alleged) theses of Leibniz's philosophy[49]. More importantly, though, Langton has been accused of having decontextualized — and thereby greatly misunderstood — the passages she used as evidence for her thesis.

Allais' approach is more cautious and more attentive to the context of claims that can be found in the Amphiboly. In her book *Manifest Reality* (2015), she presents her own well-argued metaphysical dual-aspect reading

[49] Several interpreters have discussed in detail Kant's criticism of Leibniz's monadology in the Amphiboly (see e.g. Zinkin 2008; Brandt Bolton 2021), or more generally the relationship between the two philosophers. However, that is not my aim here. I am not concerned with Kant's view of Leibniz, nor with the effectiveness and fairness of the Amphiboly's argument as a criticism of him. Also, Longuenesse's outstanding analysis of the concepts of reflection (Longuenesse 1998, ch. 6) — which predates both works I am discussing here — has broader and more ambitious aims that will not be discussed here.

— which also represents a kind of 'one-world' view[50]. According to her account, the object is one, despite having two metaphysically different sets of properties: «I argue that [Kant] thinks that things have intrinsic, non-relational, categorical natures which ground their relational appearances» (Allais 2015, p. 231). The expression 'things in themselves' therefore seems to coincide with 'the intrinsic nature of things' (or 'things in regard to their intrinsic nature'), which would in turn be the 'inner' that Kant discusses in the Amphiboly. In short, Allais claims that Kant and Leibniz agree insofar as they are «committed to something that grounds relational properties» and to the fact that we have knowledge «only of spatial (and therefore relational) properties» (Allais 2015, p. 235), while, on the other hand, the two philosophers disagree on the existence of monads. It is extremely relevant here that, in Allais' view, the fact that relational properties imply an intrinsic nature grounding them is an *analytical truth* whose negation results in a contradiction: we are thus compelled to acknowledge the existence of an intrinsic nature of things as they are in themselves. However, Allais readily acknowledges that we can never find intrinsic properties[51] in phenomenal reality (the only substance in experience being *substantia phaenomenon*), which seems to contradict the aforementioned analytical truth. Allais relies on her metaphysical dual-aspect reading to resolve this tension. She argues that Kant's solution lies in his transcendental idealism: *as appearances* things have nothing but relational properties, while *in themselves* things have an intrinsic nature that grounds such properties. Despite the impossibility of our cognizing such an intrinsic nature, according to Allais we can and indeed *must* establish its reality. That leads her to claim that Kant's distinction in the Amphiboly between what is thought by pure understanding and what is an

---

[50] In her book, Allais does not use the label 'dual-aspect view' to describe her position, but I agree with Quarfood's use of the expression to describe Allais' stance on the problem of Kant's transcendental idealism (Quarfood 2004). As for the label 'one-world', mine is of course a simplification. For Allais does not argue for a kind of one-to-one correspondence between appearances and things in themselves: a single thing in itself could in fact appear as different spatiotemporal objects. Nevertheless, what matters here is that appearances and things in themselves are *not* just completely separate and independent objects that have only some kind of unspecified relation to each other.

[51] Intrinsic properties which, as Allais correctly emphasizes, are not required by Kant even for fruitful scientific research: it is possible to advance in science while having cognitive access to relational properties only.

object of sensible intuition and cognition somehow maps onto the distinction between things in themselves and appearances.

Langton's and Allais' works have of course already been criticized to some extent[52]. Nevertheless, as far as concerns the conclusions the two scholars draw from their analysis of the inner and the outer in the Amphiboly, I think that a more textually grounded critique is needed, one that is based upon a careful examination of the Amphiboly's text. In the following, I want to argue that in the Amphiboly Kant holds that we cannot even state with certainty that the objects of possible experience have an intrinsic nature. In fact, such a statement would have no positive ontological meaning: it cannot inform us about a putative reality in itself nor about the nature of appearances.

2.2 Kant's Overall Argumentative Strategy in the Amphiboly and its Application to the Inner/Outer Dichotomy

I will now illustrate the overall argument of the Amphiboly and discuss whether Allais' reading of the inner/outer dichotomy is consistent with the text. According to Kant, transcendental reflection is the capacity to compare different representations while taking into account the cognitive faculty in which each representation is situated. He lists four pairs of concepts of reflection, which express the relation «in which the concepts in a state of mind can belong to each other». In this respect, the relation between two representations can be of «**identity** and **difference**, of **agreement** and **opposition**, of the **inner** and the **outer**, and finally of the **determinable** and the **determination** (matter and form)» (A261/B317). Overall, Kant aims at refuting Leibniz by showing how such relations lack a univocal meaning applicable to «objects in general». Rather, it is essential to distinguish

[52] Besides the above-mentioned critiques of Langton, for a critique of Allais see also Anderson (2016), Schulting (2017b) and Heidemann (2021). Schulting actually focuses on defending the 'phenomenalist' features of his own interpretation of Kant's transcendental idealism. Heidemann's article about things in themselves shows, only very briefly, a set of seriously problematic aspects of Allais' metaphysical dual-aspect view, and he does not mean to provide a strong textual analysis of the Amphiboly nor of other parts of the *Critique*. The scope of both papers is thus different from mine. As for Anderson's review, see a few notes below.

whether the representations involved are mere concepts of the understanding or appearances. In short, Kant accuses Leibniz of having considered these four pairs of relations, i.e. the concepts of reflection, only with respect to the representations of pure understanding taken as noumena (i.e. as having a non-sensible objective referent). Accordingly, argues Kant, Leibniz drew principles which he mistakenly took as valid for all (logically) possible objects in general. Kant then criticizes him by pointing to the spatiality of outer appearances, a feature that forces us to rethink the meaning of the concepts of reflection and nullifies Leibniz's allegedly analytical principles. Therefore, the conclusion is that, in order to compare representations, it is essential to take into account the cognitive faculty to which they belong.

This is what Kant does in the Amphiboly. On the other hand, when discussing each pair of concepts of reflection, I will also highlight what I believe Kant does not do. I would like to make explicit in advance the kind of strategy that I think should not be attributed to Kant. In the Amphiboly, Kant is *not* concerned with establishing analytical principles that would be allegedly valid for 'things in general', as Leibniz maintains to have done. Much less does he attempt to validate similar principles by means of the application of one of each pair of concepts of reflection to objects taken as appearances, and the other concept of each pair to the same objects, only this time taken as things in themselves. Yet this is precisely the strategy that Allais claims Kant uses with regard to the concepts of inner and outer. My aim is to show that Langton's and Allais' reading is not tenable, under the assumption that it is reasonable to expect an overall coherence in the strategy executed by Kant in the Amphiboly.

I will start by discussing the concepts of *identity* and *difference*. In line with his preliminary statements, Kant argues that the criteria for establishing diversity and identity between two representations change depending on whether the latter are appearances or objects of pure understanding — though I will show that it is one of Kant's main claims in the Amphiboly that, as I have previously stressed, the very notion of an actual 'being of pure understanding' is already itself problematic. We have already made clear that, according to Leibniz, every concept, when completely determined through conceptual analysis, analytically refers to an individual. Therefore

two individuals must necessarily be the referents of two such distinguished concepts. Conceptual analysis is Leibniz's criterion for indiscernibility, for there cannot be two referents for the same (thoroughly determined) concept. Kant argues instead that two objects (which for him are phenomena) can have the same properties, as far as concerns their concepts, and still be in two different locations in space, which means that they indeed are two *different* objects. Space is hence a condition of multiplicity and singularity. What matters for our purposes here is that Kant's claim entails his broader criticism of Leibniz, namely, that the latter did not realize the specificity of sensibility, its irreducibility to mere logical concepts and the indispensable role it plays in the formation of cognitions. This criticism of Leibniz's epistemology extends to his metaphysics as well. Leibniz was unable to understand the ontological extension of his own principles: objects of pure thought are not all objects in general, therefore the principle of the identity of indiscernibles does not hold true for all possible objects. Nevertheless, there is no reason to believe that for Kant this principle would be valid if applied to things in themselves[53]. Rather, Kant accuses Leibniz of taking objects of pure thought as positively determinable noumena, claiming thereby to determine all objects in general through mere thought. In the last section I will discuss the notion of noumenon that emerges from the Amphiboly. What is important now is that Kant clearly does *not* ascribe one element in the identity/difference pair to appearances and the other to things in themselves: it is unclear what this ascription would even amount to.

The same argumentative structure returns when Kant examines the «agreement/opposition» dichotomy. Kant *apparently* equates objects of the pure understanding and things in themselves: «If reality is represented only through the pure understanding *(realitas noumenon*), then no opposition between realities can be thought» (A264/B320). The reader might think that Kant is stating that objects of pure understanding are actually things in themselves, and that we are able to affirm that there can be no real opposition between them. Nevertheless, despite the ambiguity of Kant's wording, this reading is incorrect. A few pages later, in fact, Kant reminds the reader that pure understanding alone cannot relate to objects, and that

[53] Contrary to what Wood, for example, believes. I will discuss this issue more in detail at 2.4.

indeed there is no object of pure intellect at all: «so would it in fact have to be *if* the pure understanding *could* be related to objects immediately» (A267/B323, emphasis mine); and again several pages later, in the Remark to the Amphiboly, we find a passage that nicely conveys what I have previously argued in detail: «we cannot thereby positively expand the field of the objects of our thinking beyond the conditions of our sensibility, and assume beyond appearances objects of pure thinking, i.e., noumena, since those do not have any positive significance that can be given» (A287/B343). Therefore, when Kant uses the expression «*realitas noumenon*» he is doing nothing but adopting Leibniz's view as a hypothesis. As is often the case, Kant delves so much into the theses he is criticizing that it is sometimes unclear which thesis he is endorsing and which he opposes. As to the concepts of agreement and opposition, Kant argues that, while an opposition between Leibnizian objects of pure understanding is impossible, a *real* opposition is possible between actual objects belonging to phenomenal spatiotemporal reality[54]. The effects of two forces acting in space, when opposing each other, might suppress each other. Also in this case, as with the previous pair of concepts, the agreement/opposition dichotomy as a whole acquires a new meaning when employed to compare appearances.

Moving now to the fourth pair of concepts, namely *matter* and *form*, the argumentative strategy adopted by Kant is by now predictable. In short, when comparing Leibniz's putative objects of pure understanding, *matter*, i.e. «the determinable in general» (A266/B322), *precedes form*, i.e. determination of the determinable. However, in the case of Kant's phenomenal realm, the forms of intuition (space and time) «precede all appearances and all data of appearances, and instead first make the latter possible» (A267/B323), that is, *form precedes matter*[55]. Once again, it should be noted here that the use of the concepts of reflection changes when the objects compared are indeed appearances.

Now that we have seen how the overall argumentative structure of the Amphiboly works (and how it does *not*), I shall try to test my interpretation

---

[54] For more on logical and real opposition in Kant, see Giovanelli (2011, especially ch. 2).

[55] Longuenesse (1998, ch. 5) brilliantly shows how the priority of form actually extends to all aspects of Kant's philosophy, and how important is this fact.

more closely on the third pair of concepts, namely, *inner* and *outer*. As I previously suggested with regard to the other paired concepts of refection, these two concepts as well have different meaning depending on whether they relate to appearances or putative objects of pure understanding. Some passages might seem ambiguous at first — for example: «In an object of the pure understanding only that is internal that has no relation […] to anything that is different from it» (A265/B321). We have already seen, though, that with similar statements Kant does not mean to positively determine the object of pure understanding taken as a noumenon, since in the passage from A287/B343 previously quoted Kant writes that it is not even possible to allow for «objects of pure thinking». It is thus clear that the statement in question, however true it may be for Leibniz's concepts, has no actual objective validity for Kant. The text continues: «The inner determinations of a *substantia phaenomenon* in space, on the contrary, are nothing but relations, and it is itself entirely a sum total of mere relations». In the realm of spatiotemporal appearances the inner is reducible to the outer: the determinations of spatial objects are ultimately nothing but relations[56]. While Leibniz might conceive of representations of pure understanding as having an outer and inner nature, we learn from *experience* though that the inner of phenomenal objects is merely *relative* and not *absolute*, as their inner is ultimately nothing but their outer, i.e. a sum of relations[57]. Once again in the Amphiboly, Kant shows how the meaning of a pair of concepts of comparison — this time the concepts of inner and outer — depends on the

---

[56] Langton (1998, p. 53) correctly observes that the notion of *substantia phaenomenon* echoes Baumgarten's definition of *phaenomena substantiata* (*Metaphysica*, §193), something that appears to be a substance and yet it is not. This something is actually an accident, whose relative subsistence leads us to take it for a substance. Now, that for Baumgarten it is a kind of false substance which actually presupposes a proper substance (i.e. the monad) is indeed correct: every proper substance is a monad, for otherwise it would be nothing but a *phaenomenon substantiatum* (*Metaphysica*, §234). Nevertheless, unlike what Langton seems to believe, this fact does not prove that for Kant as well a *substantia phaenomenon* necessarily entails the existence of a *substantia noumenon*. See below in Chapter 4 (sections 4.5 and 4.6) for a discussion of the notion of *substantiale*, namely the ultimate subject of inherence taken in abstraction from its accidents. In the present context it suffices to insist that the schematized category of substance is for us the only one that has an objective content, and that we are not theoretically entitled to assert the existence of a *substantia noumenon*.

[57] It seems that Kant nonetheless allows for us to call certain determinations of appearances 'inner' in a merely comparative sense — «comparatively internal» (A285/B341).

kind of objects they relate to: objects of pure thinking or actual objects in space and time.

From what has been said thus far, the reader has no reason to expect — as Langton and Allais do — that Kant would claim that the 'inner' of objects is their in-itself aspect as a putative real object of pure understanding, and that their 'outer' is instead their relational aspect equivalent to the phenomenal manifestation of those same objects. Of course, a few lines below, Kant writes that, as an «object of the pure understanding […] every substance must have inner determinations and forces that pertain to its inner reality» (A265/B321). Nevertheless, these lines should again be read in light of A287/B343, where Kant warns us that such objects have no positive significance, for they are immediately followed by a description of Leibniz's view — more specifically, by an account of how Leibniz derives the concept of monad. Kant's concern here is that Leibniz mistakenly believes that relations among *concepts* of understanding alone are objective relations and determinations which should hold for all *objects* in general (where the latter are taken to be entirely independent of human sensibility). There is therefore no reason to read this paragraph as if Kant were claiming that relational properties of appearances necessarily entail an existing inner nature actually grounding them.

2.3 Relational Properties Without Intrinsic Natures: A Logical Impossibility?

I would now like to examine the issue from a slightly different perspective. According to Langton's and Allais' interpretations, the Amphiboly section shows that we are able to make some claims regarding things in themselves by means of pure thinking. Specifically, their thesis is that we know for sure that objects have an 'inner', an intrinsic nature, which is equivalent to the 'in itself' aspect of appearances. Allais claims that Kant holds the following statement to be logically valid: 'there are no relational properties without an intrinsic nature underlying them'. Since in the phenomenal realm there can be nothing absolutely 'inner', we must then acknowledge that the inner of things is their in-itself aspect. At this point, Allais anticipates a possible objection to the view she ascribes to Kant. As

Allais herself acknowledges, one might accuse her of violating Kant's distinction between logical and real possibility: intrinsic natures are *logically* possible, but we cannot know whether they *really* are possible. However, Allais' response to this criticism goes as follows:

> Crucially, something's being logically possible is not enough (Kant thinks) to show that it is really possible, but something's being logically impossible is enough to show that it is really impossible. […] In the Amphiboly, Kant clearly thinks that it is a logical truth that relations presuppose something non-relational. The idea of there being relations without there being something non-relational simply does not (Kant thinks) make sense; it is not any kind of possibility (Allais 2015, p. 240).

Therefore, through this analytical truth Kant would be able after all to claim something about things in themselves and their relation to the realm of appearances, namely, that they have intrinsic properties. More than that, Allais maintains that Kant is led to make this claim about the nature of things in themselves by the need to overcome the incompatibility between the logical necessity of intrinsic natures and the fact that no such nature can be found in phenomenal reality. However, in the Amphiboly Kant repeatedly deems as invalid any attempt to make claims concerning representations in general that is not based on a preliminarily *reflection* on the source of the latter. I believe that Allais' interpretation is not correct, for her reading eludes precisely this preliminary reflection — as I will now proceed to argue.

The alleged logical impossibility to which Allais refers can be formulated in the following way (which, for a reason that will be soon evident, I call 'i3'):

(i3) It is impossible for there to be relational properties without intrinsic natures grounding them.

From this proposition, following Allais, it could be said that Kant obtains a truth about things in themselves, namely, that they constitute the actual intrinsic nature that grounds phenomenal relational properties. I suggest now that we could analogously put forward three propositions concerning the remaining pairs of concepts of reflection. These too result from the comparison between representations of pure understanding in abstraction

from sensibility. If we call the previous alleged logical impossibility 'i3', then we have:

(i1) It is impossible for two things having the same internal determinations not to be identical with each other.
(i2) It is impossible for there to be any real opposition between two things.
(i4) It is impossible for form to precede matter.

Had Kant adopted the strategy attributed to him by Allais regarding the inner/outer dichotomy to obtain a *truth* about things in themselves, he could have consistently derived these other three propositions, from which we would obtain as many truths about things considered in themselves. However, as I illustrated above, what Kant actually does in the Amphiboly is rather to *refute* those very claims through transcendental reflection on the sources of representations. *Sensibility* is what makes the difference here: since sensibility cannot be reduced to conceptual features (such as Leibniz's confused perceptions), the *objective* validity of any universal claim would require taking into account the specific role of sensibility in cognition. In fact Leibniz claims to obtain truths that are valid for all objects (and therefore for what Kant calls things in themselves) by deriving them from the aforementioned impossibilities, which in turn arose from the mere comparison between concepts without taking into account the specific contribution of sensibility in the constitution of objectivity.

Here are two passages, from the Remark to the Amphiboly, that make Kant's argument explicit:

> [C]ategories alone are not sufficient for the cognition of things in themselves, and without the data of sensibility they would be merely subjective forms of the unity of the understanding, but without any object. […] And one cannot call the noumenon such an object […] The concept of the noumenon is therefore not the concept of an object, but rather the problem, unavoidably connected with the limitation of our sensibility, of whether there may not be objects entirely exempt from the intuition of our sensibility […] (A287/ B343-4).

> [S]ince something is contained in the intuition that does not lie at all in the mere concept of a thing in general, and this yields the substratum that cannot be cognized through mere concepts, namely a space that, along with every thing that it contains, consists of purely formal or also real relations, I cannot say that since without something absolutely inner no thing can be represented **through mere concepts**, there is also nothing outer that does not have something absolutely internal as its ground *in the things themselves that are contained under these concepts* and in **their intuition** (A284/B340, italics mine).

The four propositions i1, i2, i3, i4 are not genuinely valid, and indeed, while interchangeable with four other propositions which would have to be, given Allais' approach, analytically necessary, they are in particular not analytically necessary. For example, in the case that concerns us the most here: it is not truly necessary at all for relational properties to always be grounded in some intrinsic nature; therefore it is also not true that it is impossible for there to be relational properties without intrinsic natures grounding them. The phenomenal realm is not a kind of exception to the rule that Kant needs to explain by means of his idealism, as Allais believes, but it is rather the proof that those purported necessities are not logical truths in the first place[58]. This point has been missed by Allais because she has evaded the requirement of preliminary transcendental reflection mentioned by Kant at the beginning of the Amphiboly, which means that the source of the representations that are being compared needs to be taken into account. The comparison between representations of pure understanding in complete abstraction from

[58] This point is stressed by Kant in a passage from the *WRP*: «Leibniz's law of indiscernibility […] is violently in conflict with reason, since there is no understanding why a drop of water at one place should prevent an identical drop from being encountered at another. But *this collision proves* at once that, in order to be known, things in space must be represented, not merely through concepts of the understanding, as things-in-themselves, but also in accordance with their sensory intuition as appearances» (20: 280, emphasis mine).

sensibility in no way informs us about *objects* «in general»[59], either about appearances or — given that noumena, as Kant insists, are not even objects in the strict sense (cf. A287/B343-4) — about noumena. Accordingly, Kant in fact writes in a note about the relation of real opposition:

> If one wanted to make use of the usual escape here, that at least *realitates noumena* cannot act in opposition to each other, one would still have to introduce an example of such pure and non-sensible reality in order to understand whether such a reality represents something or nothing at all. But no example can be derived from anywhere except experience, which never offers more than *phaenomena* (A282-3n/B338-9n).

In the same way, objects in space are nothing but mere relations, and there is nothing wrong with that. As Kant emphasizes, they are indeed mere appearances. *If* anything objective, an individual intelligible being, corresponded to what we represent «through mere concepts», then we could perhaps say that such an object would have an intrinsic nature. However, such an object is not given to us, i.e. we are not able, as Kant demands in the note, to «introduce an example of such pure and non-sensible reality». Appearances are all we have access to; hence, the previous hypothetical statement is empty, which means that the notion of intrinsic nature has no objective application.

---

[59] This point is also stressed by Anderson's (2016, pp. 1666-7) criticism of Allais and I agree with most of his insightful observations. Nevertheless, I believe that my argument goes beyond Anderson's remarks, for it shows in detail how Allais' line of thought unnecessarily turns the Amphiboly into an incoherent argument (by generating the false claims i1, i2 and i4). And this was possible through a more extended analysis of Kant's text, that is, only after taking into account all the dichotomies (2.2) and showing concretely Kant's strategy. Allais (2016) attempts to address Anderson's remarks by claiming once more that Kant's point is merely to stress the transcendental ideality of space and the need for another level of reality, that is, the realm of intrinsic natures. However, it should be clear that this kind of reply cannot affect the argument that I propose here.

2.4 Kant's Agnosticism About Things in Themselves in the Amphiboly and the Problem of Identity

I shall now add a few concluding observations on the conception of thing in itself that emerges from the Amphiboly (and the Remark attached to it). The Amphiboly section has been used by Langton as the main source of her strong metaphysical reading of Kant, and Allais draws similar conclusions from it. In opposition to that, I believe that not only does the Amphiboly not easily fit their metaphysical dual-aspect reading, as they maintain, but on the contrary it is rather an obstacle to such a reading. In this section of the *Critique* Kant is committed to a strong form of agnosticism about the nature of things in themselves — which is perfectly in line with the considerations I made in the previous chapter. First of all, as we have seen, this is the main point of his criticism of Leibniz, which leads Kant to argue for the impossibility of formulating principles that have any objective validity on the basis of a comparison between «mere concepts». Secondly, such agnosticism is explicit in several statements — most of which have already been quoted — in the concluding pages of the Remark to the Amphiboly. Here is another example, in which, more specifically, Kant debunks the hypothesis that pure, unschematized categories per se entail a reference to some presumed intelligible object:

> If by merely intelligible objects we understand those things that are thought through pure categories, without any schema of sensibility, then things of this sort are impossible. For the condition of the objective use of all our concepts of understanding is merely the manner of our sensible intuition, through which objects are given to us (A286/B342).

Categories alone, then, «without the data of sensibility […] would be merely subjective forms of the unity of the understanding, but without any object» (A287/B343-4); that is, as I have already shown, they express the way the cognitive subject necessarily thinks of objects instead of the objective determination of a putative non-phenomenal object.

In an extremely confusing and interpretatively problematic passage[60], Kant spends a few words on intelligible beings understood as «objects of a non-sensible intuition, of which our categories are certainly not valid» (A286/B342). Unfortunately, as noted earlier (see section 1.4), Kant's language seems here inconsistent with that of the PN chapter, for there he refers to 'objects of a non-sensible intuition' as noumena taken in the positive sense (see B307), while in the passage at issue here he calls them «*noumena* in this merely negative sense», which «must of course be allowed». As I said, I find it more reasonable to stick with the terminology introduced in the PN chapter and to read the whole passage light of it. The empty thought of an object of non-sensible intuition is not only logically possible, but also useful for contrastive purposes. However, since we are not acquainted with such type of intuition, we cannot take the intuitive understanding and its purported object — for which no actual example can be introduced — as *real* possibilities. Consequently, no positive use of the notion of noumenon is permitted, i.e. we have no theoretical grounds for claiming that a certain concept (e.g. that of 'intrinsic nature') has one or more non-sensible objective referents.

Besides Langton' and Allais', another interpreter that misunderstands Kant's point about things in themselves in the Amphiboly is Allen Wood. On his account, Kant considers the identity of indiscernibles as a valid positive

---

[60] «Indeed, even if one would assume another sort of intuition than this our sensible one, our functions for thinking would still be without any significance in regard to it. If we understand thereby only objects of a non-sensible intuition, of which our categories are certainly not valid, and of which we can therefore never have any cognition at all (neither intuition nor concept), then *noumena* in this merely negative sense must of course be allowed: for they would then not say anything but that our manner of intuition does not pertain to all things, but only to objects of our senses, consequently that their objective validity is bounded, and room thus remains for some other sort of intuition and therefore also for things as its objects. But in that case the concept of a *noumenon* is problematic, i.e., the representation of a thing of which we can say neither that it is possible nor that it is impossible, since we are acquainted with no sort of intuition other than our own sensible one and no other sort of concepts than the categories, neither of which, however, is suited to an extrasensible object. Hence we cannot thereby positively expand the field of the objects of our thinking beyond the conditions of our sensibility, and assume beyond appearances objects of pure thinking, i.e., *noumena*, since those do not have any positive significance that can be given» (A286-7/B342-3).

principle that we are able to apply to things in themselves[61]. More precisely, according to Wood's so-called «identity interpretation», objects of pure understanding mentioned by Kant are obtained through the process of thinking particular appearances in abstraction from their sensible features:

> Although things in themselves cannot be sensed, appearances can be thought through the pure understanding, simply by thinking of them in abstraction from the ways they can appear to us. Thus while the sensible criterion for identity cannot apply across the gulf separating phenomena from noumena, the intelligible criterion can apply. This seems in fact to be precisely the way in which Kant himself often arrives at the concept of a noumenon or thing in itself. We begin with particular sensible things (appearances) and then represent them as they are apart from our sensing them, solely through concepts of understanding (Wood 2005, p. 69).

Moreover, since for Wood they are the same object (though considered in two different ways), we are perfectly able to compare appearances and things in themselves in reflection. Therefore, we can affirm their identity. In short, Wood's reading faces at least the following three problems. First, as I showed earlier, nothing in Kant's argument suggests that he accepts Leibniz's principle of identity as valid for things in themselves. That is not the point of Kant's argument in the Amphiboly, which instead criticizes Leibniz much more radically. This makes it questionable to speak of a useful «intelligible criterion» for identity. Secondly, we are able to compare appearances and things in themselves and state their identity, as Wood claims, only because, in the initial process of abstraction as he characterizes it, we have first already assumed that they are the same thing — which is however what has yet to be proved. Finally, indeed, Wood's claim that we are able to grasp the in-itself aspect of a particular appearance through pure understanding by a process of abstraction is at odds with Kant's rejection of the Leibnizian idea of an object in general thought by means of pure understanding alone, since «objects of pure thinking […] do not have any positive significance that can be given» (A287/B343).

---

[61] Allais (2015, p. 73) seems instead to deny the validity of the principle of indiscernibles as a metaphysical truth about noumena.

It is worth spending a few more words on Wood's interpretation of things in themselves. Wood claims:

> the criterion of identity or distinctness involving things in themselves (if we need one at all) is Leibnizian (sameness or distinctness of concept). Once we have abstracted from the sensible — e.g. the spatio-temporal — properties of the object as appearance, hence from our empirical cognition of it, it must be the same as *itself* thought solely through pure concepts of the understanding and distinct from any thing which is represented as other than it (for instance, from a *different* appearance regarded as it is in itself) (Wood 2005, p. 69).

We have seen that there is no point for Kant in counting different referents of the same concept if these referents are taken to be mere thought entities. It is then clear that it makes no sense to discuss the identity or difference of putative thought-objects, since we are aware of no criteria for their numerical distinction. Wood's notion of a thing in itself is that of a singular thought-entity with no spatiotemporal features. But given Kant's account of concept reference, we are not really able to *positively* think such purported entities as singular, that is, we are not able to connect a referent to the concept of a particular non-spatiotemporal entity. My main point is that by holding Leibniz's principle of the identity of indiscernibles as an «intelligible criterion» valid for things in themselves, and indeed already by making claims concerning the identity or difference (both of which presuppose distinctness) of things in themselves taken as entities of pure thought, Wood is making illicit, positive claims about noumena. Therefore, on the basis of the Amphiboly (and of what we have seen in the previous chapter), there is good reason to reject Wood's own brand of 'one-object' interpretation of Kant's idealism.

Wood seems committed to a very strict identity, or sameness, between an object as appearance and as a thing in itself: «If I arrive at the concept of the chair in the corner first by cognizing it empirically and then by abstracting from those conditions of cognition, so that I think of it existing in itself outside those conditions, then it is obvious that I am thinking of the same

object, not of two different objects» (Wood 2005, pp. 69-70)[62]. On this account we would be able to talk about a *chair* as it appears and the very same chair as it is in itself, in abstraction from sensibility. However, it is fair to raise an objection, at least on philosophical grounds. As Lance Hickey rhetorically asks, imagine that we «abstract away all the empirical properties of the object, i.e., all the properties that give the cup [or Wood's chair] the character of an appearance in space and time. What are we left with? Certainly not the cup» (Hickey 2001, p. 129), nor the *chair*, as Wood seems to believe. In order for us to identify an individual appearance we need at least to subsume it under schematized categories and empirical concepts, and make reference to its spatiotemporal location(s). All this is not available for us when we abstract from sensibility. Therefore, along with the difficulty of making reference to individual non-sensible beings, I also see no available criteria to establish any supposed 'cross-worlds' identity among things as appearances and as they would be in themselves.

To be sure, the above remarks on Wood raise the problem of accounting for the fact that Kant at times does mention the possibility of thinking of noumena as 'the very same' [«*eben dieselbe*» (B306)[63]] objects given to us in experience yet in abstraction from its sensible constitution[64]. Yet, in light in the aforementioned difficulties it is not clear at all to me what would exactly remain that could be in principle re-identified. Stang (2022) is sympathetic to the objections often moved against 'Identity readings' (see his 'Introduction' to the paper). Nevertheless he has a promising suggestion to make room for a sense in which there can be some substantial identity among appearances and things in themselves. It is worth taking a look at his proposal.

---

[62] See also Roche (2010, p. 677) for a similar claim (earlier discussed in note 42).

[63] As Rosefeldt notes (2022, p. 24 fn. 13), the Guyer-Wood translation is misleading here, since it translates '*eben dieselbe*' as 'other objects'.

[64] This fact is stressed, among others, by Willaschek 2001 and Rosefeldt 2022. Unlike the latter, though, Willaschek cannot be counted among the proponents of a full-fledged 'one-world' or 'dual-aspect' view. For his aim in the article is rather to argue that the notions of '*Erscheinung*' and '*noumenon*' have more than one meaning in Kant, and that the correctness of the identity claim changes accordingly. For further textual evidence of the identity between appearances and things in themselves, see for example A38/B55 and Prol. 4: 289. These passages are nonetheless counterbalanced by others leaning in the opposite direction, e.g. A165/B206 and A276/B332.

Building upon a distinction between 'object' and 'thing', Stang argues that appearances and things in themselves are always different 'objects' but can be the same 'thing'. His complex reconstruction goes roughly as follows. The concept <object> is for Kant the most general concept of transcendental philosophy, and it means in fact 'object of a representational capacity'. In order for us to think of objects as an instances of some concept thus working as objective referents in a judgement, we must think them as objects of some possible intuition. Besides our sensible type of intuition, Kant of course conceives of an intellectual one, if only as a mere logical possibility. However, given the identity statements 'x=y' (or 'x≠y'), we lack the concept of an intellect which could assign a different kind of object, namely a sensible and a non-sensible one (i.e. a thing in itself), as a value for each variable respectively. For we «have no concept of a third kind of intellect that intuits both kinds of objects» (p. 315), which are therefore non-identical in the sense that their identity is unthinkable. On the other hand, 'thinghood' is 'reality', and Kant demonstrates that 'things' are «substantial loci of intensively gradable casual force, *realitates*» (pp. 317-8). Appearances and things in themselves are then «collectively» the same 'things', i.e. in the same sense in which some goblets can be the same thing as the mass of silver that was extracted from the mountains to make them. Establishing that certain appearances are «distributively» the same thing as certain things in themselves is more complicated though. Nevertheless, Stang illustrates a special case in which this occurs. Through my awareness of the moral law I am aware of being a noumenal 'thing' endowed with noumenal freedom; yet I am at the same time aware of myself as a phenomenal 'thing' subject to deterministic causality, both of which are «one and the same locus of capacity and act» (p. 331). It is not entirely clear to me, though, why on Stang's account it would not follow that my noumenal and empirical selves are the same 'object' as well. Let us leave aside his claims on the self, though.

I would like to draw the attention instead on his claim that an appearance can be the same 'thing' as a thing in itself in case both are the «substantial loci of intensively gradable casual force». As we have seen in the previous chapter, nothing prevents us from employing, in mere thinking, the categories of 'substance' and 'causality' involved in Stang's claim in abstraction from sensibility — although the thought would thus be empty,

i.e. we cannot establish whether it has or even could have an objective referent. However, there is something more problematic with Stang's formulation. His description of such force as «intensively gradable» brings us immediately to the second of Kant's principles, namely the Anticipation of Perception. As I have stressed though (section 1.5), all principles are constitutively valid for phenomenal reality alone, and it is thus at the very least questionable whether a thing in itself can be said to have an intensive magnitude. But let us take a look at the principle:

[In the first edition:]

> The **principle**, which anticipates all perceptions, as such, runs thus: In all appearances the sensation, and the **real**, which corresponds to it in the object (*realitas phaenomenon*), has an **intensive magnitude**, i.e., a degree.

[In the second edition:]

> <Its principle is: **In all appearances the real, which is an object of the sensation, has intensive magnitude**, i.e. a degree.>

To be sure, accounting for this principle is made especially difficult — in particular in the second edition — by Kant's talk of an 'object of the sensation'. For one may read the principle as a claim about the thing in itself that affects our sensibility. This would imply that the thing in itself be a *realitas* that has an intensive magnitude. In any case, this is by no means the only possible reading. Indeed, the opposite seems more plausible.

In his thorough analysis of the Anticipation, Marco Giovanelli correctly stresses the importance of Kant's reference here to «*realitas phaenomenon*» (Giovanelli 2011, in particular ch. 1 and 2). Establishing that *realitas phaenomenon* has an intensive magnitude leads Kant to overcome the traditional way of conceiving opposition and negation as merely logical. As we have previously seen, whereas logical opposition — i.e. contradiction leading to nullification of the opposites — is the only way to conceive opposition among purely intelligible things (*realitas noumenon*), *real* opposition and negation are of a whole different sort; and this is a

consequence of Kant's principle that *realitas*, in the phenomenal world, is a matter of degree. Now, here is where Stang's insightful account turns out nonetheless unclear and problematic. Stang acknowledges that the Anticipation, combined with the first and second analogies, proves that «objects of *experience* necessarily possess an intensive magnitude of causal force, *realitas*» (p. 318, first emphasis mine). From here, though, in the space of a page, Stang arrives at the conclusion that it is possible that «appearances and things in themselves are distinct *objects* but the same *things*» (p. 319). He does so by appealing to arguments concerning the independent existence of the objects that must affect our organs in order for representations in us to obtain, which we then synthetically combine. It is not my aim in this chapter to address the problem of affection[65]. What matters here is that no convincing argument is provided by Stang in support of the claim that characterizing things in themselves as «substantial loci of intensively gradable casual force» is compatible with the fact that the 'real' that we know a priori to have an intensive magnitude is precisely only *realitas phaenomenon*. It seems to me that Stang ignores the danger that would stem from his ascription of an intensive magnitude to *realitas noumenon* as well, because this runs the risk of making Kant's claims on real versus logical opposition pointless, and is oblivious to the limits of the objective validity of Kant's synthetic a priori principles.

There is at least another promising candidate that may be characterized as that of which we would say that it is both a thing 'in itself' *and* also an 'appearance'; and it may thus be characterized as what remains the same when we abstract from sensibility. It is the *transcendental object*, a notion which is employed by Kant in a disorienting variety of ways and which, I believe, is symptomatic like no other of the intrinsic ambiguities of Kant's

[65] See Chapter 5 below.

transcendental idealism[66]. Rosefeldt (2022), for example, defends an ontological dual-aspect view and takes the notion of transcendental object to be precisely that 'something = x' of which «we know nothing […] but the way it appears to us in sensible intuition» (p. 23). Let us assume that Kant does actually employ the notion of transcendental object to indicate 'that' which supposedly has both a sensible and a non-sensible aspect. Yet, I now argue, this would be of little help for our present discussion. If we consider what can be fairly called an empirical object, e.g. the water bottle in front of me, we could say that it is an appearance (and in some sense a representation) of 'something = x' which has a constitution in itself, i.e. which is in itself non-sensible. And if there were multiple non-sensible entities that appear to us in space and time as the individual water bottle in front of me — leaving aside the unsolvable difficulties about counting noumenal entities — we could then simply call 'x' the very collection of such entities. The main problem is that it is still logically possible that different appearances which we call objects of experience correspond to the same 'something = x'. At this point one could apply the previously mentioned strategy used by Stang and

---

[66] See Willaschek (1998), which nicely summarizes the ambiguity of this notion as follows: «Wir müssen also drei Bedeutungen von „transzendentaler Gegenstand" unterscheiden: (1) die durch die reinen (unschematisierten) Kategorien bestimmte Struktur eines „Gegenstandes überhaupt", (2) das für alle Erscheinungen identische „Etwas = x" und (3) den für unterschiedliche Erscheinungen jeweils spezifischen, uns aber unbekannten transzendentalen „Grund"» (p. 334). For a slightly different classification of the passages, see Bird (2006, ch. 15 section 1.2). Accordingly, some scholars take the connection between the transcendental object and the thing in itself to be indicative of Kant's commitment to the existence of the latter (e.g. Onof 2011, and Rosefeldt 2022, the latter of which will be discussed in a moment), while others take this link to express precisely Kant's agnosticism towards things in themselves (this is most radically defended, for example, by Hickey 2001 and Bird 2006). To get a quick grasp of how interpreters can attempt to provide a consistent account of Kant's usage of the notion of 'transcendental object', consider Gardner's (1999) suggestion that the latter has two uses, or can be realized in two different ways. First, as «a point to which the elements of the manifold of intuition may be referred, allowing appearances to be determined as thinkable objects of intuition» (p. 99). Employed in this way, Gardner submits, it is completely independent from the concept of a thing in itself. However, he claims that Kant's assumption of the existence of things in themselves leads him to a second use of the notion of a bare 'something = x', i.e. of the transcendental object: «the transcendental object is the sole representation by means of which we can think of things in themselves» (p. 100). See also de Boer (2014 and 2020), who attempts to show how the different usages of 'transcendental object' (and related concepts) should be understood in light of the different aims pursued by Kant in the various parts of the *Critique*. To be honest, though, I do not have clear enough view on the topic to build upon, nor am I entirely convinced that Kant's usage of that concept can be made consistent.

claim that appearances and things in themselves are merely *collectively* the same: only this time not the same 'thing', as Stang has it, but rather the same 'transcendental object = x'. We should therefore perhaps speak more properly of a transcendental object of the entire matter. Now, not only it is questionable that this is what Kant has in mind when he characterizes the notion of 'transcendental object', for instance, as the «correlate of the unity of apperception» (A250). More than that, this reconstruction would still not be able to account for the aforementioned literal understanding of B306 which takes Kant to claim that, given «certain objects» — not 'all objects altogether' or 'matter in general' — these are «*eben dieselbe*»[67] as what remains when we abstract from sensible features[68]. My conclusion, however unsatisfactory it may sound, is therefore the following: in light of the difficulties that I have extensively discussed, we should not take B306 and similar passages so literally as to mean that, when abstracting from sensible properties, we know something to remain and to be still the very same thing/object.

***

To recap our discussion of the Amphiboly section, Kant claims several times there that we are not entitled to make positive statements about things in themselves. We are not able to make a *positive* use of the concept of noumenon, i.e. to establish «whether such a reality represents something or nothing at all» (A282-3n/B338-9n). At the same time, Kant explicitly writes that, taken in the *negative* sense, the concept of noumenon is a problem for us to reflect upon concerning the limitation of our faculties. As Kant emphasizes (see again A282-3n/B338-9n), we cannot — at least from a theoretical standpoint — escape from the necessity of taking into account experience and its forms, and we cannot claim objective validity for concepts that find

[67] We find the same expression, for instance, in a relevant passage of the B-Preface. There Kant writes, with reference to «mere objects of **experience**», that «even if we cannot **cognize** these same objects as things in themselves, we at least must be able to **think** them as things in themselves [*daß wir eben dieselben Gegenstände auch als Dinge an sich selbst, wenn gleich nicht **erkennen**, doch wenigstens müssen **denken** können*] (BXXVI). I take the present reservations on Rosefeldt's reading to apply to this complex passage as well.

[68] I am not hereby denying that the concept of transcendental object plays a decisive role in making sense of Kant's conception of thing in itself. My argument is rather only that the former notion cannot be deployed to secure the numerical identity between things in themselves and appearances.

no sensible correlate just by purporting to apply them to things in themselves. In the third chapter of her book, Allais insists on the distinction between the negative and positive conception of noumena, and has no problem admitting that we have no insight into the latter, that we cannot even be certain about their mere existence. On the other hand, she also notes that «Kant makes it clear that his notion of things as they are in themselves should be understood in terms of a negative conception of noumena» (Allais 2015, p. 61). However, we have just seen that, employed in this negative way, i.e. as the concept of a problem, the notion of noumenon does not entail the real possibility of actual objects at all, and thus noumena cannot be positively determined, not even minimally. Therefore, as we in fact learn from Kant himself in the Amphiboly, we have no right to claim that things in themselves have an intrinsic nature — contrary to what Allais and Langton believe. What we are left with in the Amphiboly is spatiotemporal reality and the *problem* of its bounds and grounds.

3
What I am Aware of in Pure Apperception

There surely are many candidates in Kant's philosophy that bear the name of 'self', both in the theoretical and practical domain. It is not always easy to discern the cases in which Kant is speaking of the same thing in different contexts, from those where he has in mind different entities altogether. For example: the 'I' of the 'I think'; the subject endowed with transcendental and empirical faculties; the soul polemically discussed in the Paralogisms; my empirical psychological self, and the noumenal free moral agent — are these notions supposed to refer to one and the same being? How are they interrelated? In the space of a chapter I cannot attempt to give an exhaustive answer to this question — if there is any. My more modest aim is instead to consider Kant's overall stand towards the self as it results in particular from the B-Deduction and in the Paralogisms, namely those chapters where the issue is discussed most explicitly in the first *Critique*. The main question I shall address is the following: What is the 'I' of pure apperception? And relatedly: Does Kant's theory of pure apperception entail a commitment to the existence of a noumenal self on theoretical grounds in a way that contradicts his restrictions on our access to intelligible beings? I shall argue that neither the core of the Deduction nor that of the Paralogisms directly involve a commitment to the existence of a noumenal self. I will defend a reading of pure apperception that enables us to secure its *transcendental* status, while at the same time accounting for it as an actual (and thus in some minimal sense psychological) *act* of self-awareness: pure apperception is a trascendental act in the sense that its possibility is a normative demand in order for cognition to obtain. More specifically, I will suggest that pure apperception is *not* the awareness of myself as an empirical nor as a supposedly noumenal thing, but rather of myself as no more than an instance of 'logical self'.

3.1 The Principle(s) of Apperception

Despite its central role, there is very little agreement among scholars on what pure apperception amounts to. What kind of mental state is it, if any? And most importantly for us here: What is it about? In order to tackle these issues, it is indispensable to have a closer look at §16 of the *Critique of Pure Reason* and at some of its many complications. The main goal of §16 is to introduce the principle of the synthetic unity of apperception. Now, some interpreters argue that the principle is not *one*, but rather *threefold*, or at least that it entails three distinct but interconnected claims. In the following, I will draw on the formulation of the three principles of apperception proposed by Dyck (2017)[69].

The first principle identified by Dyck is «*principle*$_C$», where 'C ' stands for 'cognition': «for any manifold of my representation, it must be possible to become conscious of it insofar as it is to figure in a cognition» (Dyck 2017, p. 37). If a representation is to be part of a cognition, I must be able to be conscious of it, and precisely as *mine*, i.e. I «*must be able*» to attach the representation «I think» to it. This principle broadly corresponds to the first lines of §16. The verb 'must' ('*muß*') suggests that Kant does not aim to provide a mere description of the empirical functioning of our mind, but is rather identifying a normative demand. The choice of the verb 'be able' ('*können*') give us the additional hint that it is not necessary for me to explicitly attach the representation 'I think' every time I am conscious of a representation in cognitive contexts[70]. It must be noted that Kant does not explicitly mention cognition at the beginning of §16. Nevertheless, I agree with Dyck that there are good reasons to believe that the conditional necessity stated by Kant is only valid if representations are to be something for me in a cognitive sense (see also Allison 2004, p. 164).

---

[69] A similar distinction is also defended by Allison (2004). It is not my aim here, though, to reconstruct and assess Allison's or Dyck's more specific views on the topic.

[70] The exact implications of the formulation '*muß… können*' are deeply controversial and have been discussed numerous times among scholars (see e.g. Heidemann 2012 and Schulting 2012). The issue is especially relevant for the ongoing debate on conceptualism/ non-conceptualism in Kant (for an overview of the debate, see Schulting 2016 and McLear 2020b).

Next is the principle of the necessary unity of apperception (NUA): «for any manifold of representations that belongs to me, it must be possible to think the identity (or unity) of the subject with respect to it» (Dyck 2017, p. 37). I believe that this second principle is best conveyed by the sentence: «the manifold representations, which are given in an intuition, would not be one and all *my* representations, if they did not all belong to one self-consciousness» (B132)[71]. The identity of the thinking I is a condition for a set of representations to be mine as *one* thought. Were each representation thought by a different self, the former would not be unified in one single thought[72]. The second principle somehow relates to what Wuerth names 'Virgil Argument'. In the *Metaphysik Herder* lecture transcripts (1762-4) we read: «If each of 100 persons knew a verse from Virgil by heart, would they therefore know all of Virgil by heart? If the latter is not so, thus does matter likewise not think» (*MH* 28: 44; cf. *ML*$_1$ 28: 266-7). This is a pre-critical argument that is explicitly mentioned and criticized by Kant himself in the Second Paralogism (A351-3)[73] — albeit it does not include the example of Virgil. Wuerth argues that there is nonetheless a strong continuity between Kant's earlier views and those expressed in the *Critique*. Indeed, on Wuerth's account, Kant would retain in the critical period a form of the argument that includes the view that the soul is in some sense given to us as a simple noumenal substance. I am willing to acknowledge that a sort of critically purified version, as it were, of the Virgil Argument plays a major positive role in the Paralogisms and indeed conveys the 'second' principle of

---

[71] I chose Kemp Smith's translation here, since Guyer and Wood translate '*einem* Selbstbewußtsein' with '*a* consciousness' instead of '*one*'. In his Italian translation, Chiodi as well choses to emphasize the unity of the consciousness to which representations must belong by translating '*einem*' with '*una sola*'.

[72] Allison states the second claim in the principle apperception as follows: «Although each of the representations that collectively constitute the single complex thought could conceivably be distributed amongst a multiplicity of thinking subjects, thought itself could not be. It must be thought by (or ascribed to) a single subject» (Allison 2004, p. 164).

[73] There Kant famously refers to it as the «Achilles of all the dialectical inferences of the pure doctrine of the soul» (A351), for it is the most powerful among the arguments of rational psychology. Yet Kant shows that it ultimately fails to prove what the rational psychologist takes it to prove.

apperception (NUA)[74]. Again, the thrust of such argument would be that a thought cannot be considered *one* in case each of its parts is thought by a different subject, i.e. in case the manifold does not belong to *one* consciousness. Yet, as I shall argue below, Wuerth's hypothesis that some minimal conclusion about my self as noumenal soul would be thereby justified is definitely ruled out by what Kant writes in the Paralogisms. A central question to be addressed is then what does it mean to think a manifold of representations as belonging to one subject, and how can this subject be characterized.

Before attempting at an answer, let us consider the third and final principle, namely that of the original synthetic unity of apperception (OSUA), which I take to refer to the following claim: «Synthetic unity of the manifold of intuitions, as given *a priori*, is thus the ground of the identity of apperception itself, which precedes *a priori* all **my** determinate thinking» (B134). The principle establishes a relation of reciprocity or co-originality between pure apperception and synthesis. For, on the one hand, the synthetic unity of the manifold results from an act of combination (synthesis) by means of which representations are ascribed to one thinking subject. Yet, on the other hand, a perhaps more striking implication of the principle is that the unity of self-consciousness is itself possible in relation to the very synthesis: we represent ourselves as one thinking subject only because a manifold of representations is ascribed to one and the same 'I'.

It is useful to distinguish between at least four interconnected elements which play together a major role in all phases of Kant's argument. (1) The unity of a *synthesized manifold*, namely the objective result of the synthetic activity that combines the manifold given in intuition. It is easy to see that securing this unity is one of Kant's main goals in the Deduction. In a nutshell, the Deduction aims to show that experience, which implies cognition of objects, involves the *objective* unity of apperception (cf. §18), through which «all of the manifold given in an intuition is united in a concept of the object» (B139). This unity is necessary, unlike the subjective unity which is governed by the rules of association and which is thus accidental (cf. B140). Given the link between judgement and synthesis,

[74] Cf. *Anth.* 7: 134n.

categories are the 'ways' in which the manifold is brought together under the objective unity of apperception. Despite its central relevance, however, I am not much concerned with the unity of the synthesized manifold here. Another element is (2) the *synthetic activity*, the act of combining the manifold. It gives rise to many questions such as: How should we conceive this act? What kind of awareness do we have of it?[75] How does it relate to judgement? Again, while I will not discuss these issues, another question deserves our attention: What is the metaphysical ground of such activity, i.e. is there a being that carries out the synthesis? My main claim for now is that the principles of apperception work independently of this metaphysical problem and do not include an answer to it — which does not amount to say that the question itself is irrelevant for Kant.

I would like to bring special attention to the last two elements: (3) the identical *subject* to which thoughts must be ascribed insofar as they are mine, namely the I that figures in the representation 'I think'; and (4) *pure apperception*, a special type of self-awareness to be better defined. Let us now have a closer look at these two elements.

3.2 Apperceptive Self-Awareness and the 'I'

The principle(s) of apperception is valid a priori for it is not a description of our empirical mental life, but rather a normative demand that sets the conditions of objective, necessary unity among representations. The unity of the subject to which representations are to be ascribed is one of the conditions that makes objectivity itself possible. I submit that the thinking subject must be conceived as the point of convergence, as it were, of those representations of which I can say that they are mine. It is the original pole to which thoughts are to be ascribed if any synthetic unity of the manifold is to be possible. In this sense, the I of the 'I think' or '*logical self*'[76], unlike what an

---

[75] I believe that pure apperception cannot be construed as strictly equivalent to the awareness of my synthetic activity. For an argument against such identification, see e.g. Rosefeldt (2006).

[76] See *WRP* 20: 270, where Kant draws a distinction between «logical» and «psychological» self.

empirical or noumenal self may be, is not an individual object. It is rather a sort of '*status*' that something may possess. With an imperfect analogy, we may conceive it similarly to the juridical notion of person. The latter as well is not an individual in the world, but rather a status which conveys a series of conditions for some rights, duties and so on to be applicable.

Since the logical self is not a proper object, it is easier to understand how it can be co-original with synthetic activity itself. It is misleading to ask whether the logical self exists before a certain act of synthesis, because it is itself constituted through that act. As I read the 'third' principle of apperception, it is only because synthesis occurs that I am able to refer to my self as an 'I', as the point of convergence to which representations are to be ascribed in order for them to be cognized as a unified whole. Now, let us assume that a concrete subject exists (be it empirical and/or noumenal) which is the being that concretely carries out the synthesis: Is it the same thing as the logical self? The answer is both yes and no. In one sense *yes*, because I, as an existing individual that carries out the act of thinking and of synthesizing a manifold, must refer to myself as the subject of my thoughts. Or, to put it differently, it must be possible for me to be aware that the conditions for objective unity are fulfilled, and thus that I do *instance* the logical I to which my representations are to be ascribed. It must be possible for me to be aware that I bear the status of 'logical self'. In a trivial but important sense, however, the answer is *no*, for I as a concretely existing being am not equivalent to a status, to a formal condition. Analogously, I may view my individual self as an instance of juridical person, and yet I am not equivalent to such abstract entity that does not exist in time (nor as a noumenal atemporal thing).

What is then pure apperception? On my account, it is precisely the just mentioned reflective awareness, which I must be able to have, that the conditions for the objective unity of the manifold to take place are indeed fulfilled; the awareness that, as long as I am thinking, the logical self is instantiated. The non-temporality of the logical self entails that apperception is not a privileged access to the existence of a being that perdures through time, or that subsists outside time (in some ungraspable noumenal sense). The act of apperception, as a real act, can be viewed as happening in time, but only because it can be carried out at *any* time. Is this compatible with the

a priori character of pure apperception? There are different possible ways to make sense of the latter feature[77]. As I see it, the purity and apriority of transcendental apperception is secured by the fact that the logical self (and the representation 'I think') is not obtained from experience, and by the fact that it is something which is necessarily possible for me to think in any act of cognition. The 'I think', as a condition, logically precedes any *specific* experience whatsoever[78]. Apperception is *occasioned* by experience, though not derived from it; and it may occur at any time, but it is not *about* a being existing in time. It is of utmost importance to appreciate that the fact that, *at any given moment in time* in which I think, I may refer to myself as one and the same 'I' to which my representations are to be ascribed does not imply that I thereby cognize my self as an identical object existing *throughout time*[79].

A legitimate and crucial question is then how pure apperception fits within Kant's taxonomy of epistemic states. Unfortunately, Kant himself seems to struggle to find a place for this peculiar kind of awareness. He claims that pure apperception entails the awareness of my existence,

---

[77] Cf. Emundts (2022), who discusses some possible options.

[78] Patricia Kitcher has a different way of making sense of the aprioriness of the 'I think' which allows her at the same time to put a special emphasis on its dependence upon experience: «*a priori* representations and principles arise through the (scrutinizing) activities of the mind […]. The tendency to scrutinize by a principle associated with the 'I-think' precedes cognition; the representation of the unity of consciousness would be produced by that scrutiny» (Kitcher 2011, p. 135). Kitcher quotes a passage from *On a Discovery* (8: 221-23) as textual support: there Kant denies the existence of innate representations, but nonetheless admits inner grounds in the subject that make it possible for the form of intuitions and the synthetic unity of the manifold to arise. If I understand her correctly, Kitcher takes these inner grounds — namely the fact that we are constituted the way we are — to secure the apriori status of a representation (and thus of the 'I think'). In my view, this is not correct: the aprioriness of the 'I think' does not merely consist in the temporal precedence of a certain psychological «tendency» with respect to any actual cognition. Rather, it expresses the fact that it is a spontaneously produced representation, whose possibility is, moreover, a necessary condition of experience.

[79] I agree with Kitcher's observation that no metaphysical claim about the identity of the *faculty* for apperception throughout time is needed here either. As she nicely puts it: «Because of the ongoing necessity of referring to a common 'I', subjects must always have a faculty for doing so. That does not imply that the faculty must be the same — whatever exactly that would mean — but only that there must always be a faculty or capacity to refer different representations to the same 'I'. […] Kant's considered position seems to be not that the power of apperception is unchanging or numerically identical, but rather that subjects must always have a power of apperception, because they must be able to refer different representations to one and the same 'I-think'» (Kitcher 2011, pp. 164-5).

although he quickly adds the enigmatic remark that the concept of *existence** at work here does not amount to the homonymous category (cf. B423n)[80]. This awareness is alternatively described as an «indeterminate perception» (B423n), as «a feeling of an existence» (*Prol.* 4: 334n), and as «a **thinking**, not an **intuiting**» (B157). It is far from obvious, though, what he means by these expressions, and in any case they seem hardly compatible with each other. Moreover, it goes without saying that the hypothesis of intellectual intuition is out of question for Kant[81]. I ultimately agree with Heidemann (2013) and Proops (2021, p. 84) in acknowledging that Kant leaves us without a definitive and consistent solution to the problem of the epistemic status of the awareness of my existence* in pure apperception. It is nonetheless possible to have a better grasp of pure apperception by negative means, namely by considering that, first of all, it should not be conflated with the inner perception of my existence in time, as Kant insistently reminds us. Secondly, the awareness of my existence* is not inferred from the fact that some synthetic activity takes place, since Kant argues instead that 'I think' and 'I exist*' are identical propositions, i.e. that 'I exist*' is analytically entailed in the 'I think' (cf. B422 and A354-5)[82].

The ambiguity about the existence* claim opens the space for a reading such as the one I am proposing here, to the effect that the existence* immediately entailed in pure apperception is nothing other than the very 'instantiation' of a condition. 'I exist*' means 'I think' because I — whatever entity I may be — instantiate a condition for the objective unity of representations, i.e. I am the identical pole of self-ascription. While it is surely an unusual meaning of the term 'existence', this is compatible with the fact that Kant himself employs the concept in an unusual way which admittedly differs from the category of existence, and with the fact that he seems to do

---

[80] Henceforth I will add an asterisk — hence: existence* — to refer to this special meaning of the term.

[81] Given this unsettling epistemic void, Fichte's claim that pure apperception represents an instance of intellectual intuition in Kant's system (and that another one is the 'fact of reason', i.e. our awareness of the moral law) does not strike me as entirely absurd (see *Zweite Einleitung in die Wissenschaftslehre*, 225). It is worth mentioning that Fichte is well aware of Kant's ban of intellectual intuition, and yet he submits that his way of conceiving this act differs from Kant's and does not go against his teachings.

[82] *Contra* Kohl (2020), as I claim below at 3.3.

so for lack of a better term. My main point here is that the principle of apperception is *not* a metaphysical claim about the nature or existence of an actual individual that synthesizes and apperceives.

To be sure, the self that synthesizes and apperceives may be taken to be an object, be it empirical, or perhaps noumenal, or even both. For instance, if we decide to address the issue within the empirical domain, we need to deal with the problems related to empirical self knowledge, e.g. mind-body relation, the scientific status of empirical psychology, the possibility of applying categories to the manifold of inner sense, and so on[83]. On the other hand, we may further question whether our inner states are ultimately grounded in a noumenal substance, and whether the positing of spontaneous activity in general requires us to establish the existence of a noumenal self. Yet all these issues do not *directly* concern the principle of apperception. For example, the objective unity of the manifold would not be affected by skepticism about personal identity: even if I were not the same entity throughout time, what the objective unity of representations demands is only that, whenever I judge and cognize an object of experience, I must be able to refer to myself as the identical logical I in all my representations. Crucially, Kant argues in the Paralogisms chapter that the fact that I see myself as the identical *subject* of my thoughts from a first-person standpoint teaches us nothing about the identity of an existing *object*. More generally:

> Thus through the analysis of the consciousness of myself in thinking in general not the least is won in regard to the cognition of myself as object. The logical exposition of thinking in general is falsely held to be a metaphysical determination of the object (B409).

The claim that pure apperception is not a metaphysical determination of the self holds throughout the Paralogisms chapter. Numerous reconstructions of the Paralogisms have been offered and it is not my aim here to contribute to the list[84], as I will limit myself to a few specific remarks. It is a matter of

---

[83] For contrasting views on empirical self-knowledge, see for example Sturm (2001), Allison (2004, ch. 10), Bader (2017), Chignell (2017) and Kraus (2019 and 2020).

[84] Among the many, see for example the seminal book from Ameriks (2000), Dyck (2014), Longuenesse (2017, ch. 5), and the recent, impressively detailed analysis provided by Proops (2021).

debate what the Paralogisms chapter attempts to prove. As it happens with the Dialectic altogether, the discussion traditionally revolves around its *pars destruens*, which, in the case at issue, can be roughly depicted as a critique of the pretensions of the so called 'rational psychology' to establish the immortality of the soul as a simple, identical, immaterial substance. There is however less consensus on what its *pars construen*s — if there is any — amounts to. One reading involves some regulative use of the idea of the soul, e.g. as the guideline of psychological self-knowledge[85]. Another option, which does not exclude the previous one, is to detect Kant's positive but scattered claims describing the features of apperceptive self-awareness, i.e. about what is entailed in the 'I think'. On such line of interpretation, we should read these claims as elements of the very «logical exposition of thinking in general» mentioned by Kant in the previously quoted passage at B409. I am indeed sympathetic with the suggestion that finds Kant illustrating the «semantic rules of self-reference» (Kraus 2020, p. 147), or the «formal features of the representation 'I'» (Rosefeldt 2017)[86]. This emerges more glaringly in the B-edition. First, Kant makes the following general claim:

> All *modi* of self-consciousness in thinking are therefore not yet themselves concepts of the understanding of objects (categories), but mere functions, which provide thought with no object at all, and hence also do not present my self as an object to be cognized (B406-7).

After that, he puts forward four analytic propositions entailed in the 'I think' — which plausibly convey four of the aforementioned «*modi* of self-consciousness» — and warns us not to equate them with four unwarranted metaphysical claims corresponding to each. For example:

> that the I that I think can always be considered as **subject**, and as something that does not depend on thinking merely as a predicate, must be valid — this is an apodictic and even an **identical proposition**; but it

[85] This is the central interpretative proposal of Kraus (2020). See also Dyck (2014, ch. 7).

[86] See also Longuenesse 2017, pp. 130-1.

> does not signify that I as **object** am for myself a self-**subsisting** being or **substance** (B407)[87].

At the end of the first chapter (see section 1.5) I noted that in the Paralogisms, in particular in the A-edition, we find claims about a «merely transcendental use of the category» (A402) and the employment of the concept of substance as a «pure intellectual concept» (A403). Kant also concedes that the proposition «**The soul is substance**» is admissible if it «signifies [*bezeichne*] a substance only in the idea but not in reality» (A351). Now, understanding the implications of similar claims in the Paralogisms turns out to be especially difficult because it is not easy to discern when Kant is making claims about either (a) the legitimate conception of I of the 'I think'; (b) the idea of the soul in its legitimate regulative use; (c) the manifold given in inner sense; or finally (d) some conflation of these elements operated by Kant's polemical target, i.e. the 'rational psychologist'. My view is that, when Kant makes positive claims concerning some pure use of the categories in the Paralogisms he is referring either to the way we form and make use the regulative idea of soul, or to their employment to display the just discussed features of apperceptive self-awareness. In the latter context, though, Kant submits that concepts such as <unity> (cf. B131) or, as we have previously seen, <existence> (cf. B422-3n) are employed in a sense that must

---

[87] Here are the remaining features of apperceptive self-awareness and the corresponding unwarranted metaphysical claims: «2) That the I of apperception, consequently in every thought, is a **single thing** that cannot be resolved into a plurality of subjects, and hence a logically simple subject, lies already in the concept of thinking, and is consequently an analytic proposition; but that does not signify that the thinking I is a simple **substance**, which would be a synthetic proposition. […] 3) The proposition of the identity of myself in everything manifold of which I am conscious is equally one lying in the concepts themselves, and hence an analytic proposition; but this identity of the subject, of which I can become conscious in every representation, does not concern the intuition of it, through which it is given as object, and thus cannot signify the identity of the person, by which would be understood the consciousness of the identity of its own substance as a thinking being in all changes of state […]. 4) [That] I distinguish my own existence, that of a thinking being, from other things outside me (to which my body also belongs) — this is equally an analytic proposition; for **other** things are those that I think of as **distinguished** from me. But I do not thereby know at all whether this consciousness of myself would even be possible without things outside me through which representations are given to me, and thus whether I could exist merely as a thinking being (without being a human being)» (B407-9). See again Kraus (2020, section 4.3.2) and Rosefeldt (2017, pp. 234-6) for different attempts of a precise reconstruction of these positive features.

be kept distinct from the one expressed by the homonymous categories. This may be taken to be justified by Kant through the striking claim that the thinking self «cognizes itself through the categories, but that it cognizes the categories, and through them all objects, in the absolute unity of apperception, and hence cognizes them through itself» (A402). Apperceptive self-awareness is thus more original then the categories themselves. Accordingly, some interpreters contend that this is precisely how the deduction of the categories works, namely from the original structure of self-awareness in apperception[88]. Overall, I take these considerations to support what previously established (section 1.5), as they seriously undermine the hypothesis that finds Kant positively determining the nature of a noumenal self through unschematized categories in the Paralogisms chapter. More than that, I hope to have shown thus far that, as expected from the results of our previous chapters, the principle of apperception and the conditions for self-ascription need not directly involve a reference to an individual noumenal self (or to whatever purported set of noumenal entities may jointly be taken to act as an individual self).

3.3 What Pure Apperception is Not

I shall now conclude my analysis by engaging with a couple of representative interpretations of pure apperception which are particularly useful for underscoring by contrast the key elements of the position I have been defending in this chapter. The latter can be resumed as following: Pure apperception is indeed an *act* that I must be able to perform at any given time; and it is a special kind of self-awareness, although *not* of myself as an empirical or noumenal *individual object* that grounds synthetic activity.

Any interpretation of pure apperception must be able to account for the pivotal distinction drawn by Kant between it and *empirical* apperception. The latter is indeed the empirical awareness of my ever-changing inner states and, through these, of myself as a being that exists in time. Empirical

[88] Schulting (2018) attempts at a systematic reconstruction of how categories are deduced from apperception. See also McLear (2020a), who focuses on the category of substance to argue that categories are grounded in our intellectual self-awareness in pure apperception.

apperception presupposes the pure one: «The empirical consciousness that accompanies different representations is by itself dispersed and without relation to the identity of the subject» (B133). What Kant has in mind when he speaks of such empirical consciousness is likely something of the following kind: for instance, I am aware of being a person named Rodrigo and having certain character traits (though others may not agree with my self-depiction), I recall (perhaps erroneously) that I felt happy yesterday and annoyed getting out of bed this morning, and at the present moment I am aware of being focused on writing on Kant. Crucially, empirical apperception cannot secure a necessary a priori connection among representations.

By contrast, Kant is eager to stress the a priori, transcendental nature of pure apperception. However, a question arises: if pure apperception is an actual *act*, what does it mean for something like an act to be *a priori*? An act of self-awareness would seem to be a psychological action of the mind occurring in time, after all. How then could this psychological feature of mental acts be compatible with the aprioriness of pure apperception? How is it possible not to conflate pure with empirical apperception?

One way to overcome these worries is to construe pure apperception not as an act of self-awareness, but rather as the mere *form* of self-consciousness in general. This strategy is pursued, for instance, by Kraus (2020)[89]. In line with her hylomorphic account of mental faculties, Kant's pure apperception is described as the «*transcendental* faculty for consciousness in general» (p. 105), the *form* of reflexivity (see Kraus 2020, section 3.4). In this way, pure apperception provides a formal condition of every representation that can be meaningful for me, and it is thus a necessary (but not sufficient) condition to represent an object of cognition. At this point another feature of Kraus' proposal comes into play, namely an «expressivist» interpretation of the 'I think'[90]. Being the mere *form* of reflexive consciousness, the original synthetic unity of apperception has no representational content at all. Some minimal

---

[89] Kraus indeed criticizes, among others, those interpreters who aim to account for pure apperception in terms of an awareness that has the logical I as its content (Kraus 2020, section 3.3). While she nonetheless holds the view, mentioned above, that in the Paralogisms chapter Kant provides «semantic rules of self-reference», she takes this to concern our «Logical self-consciousness» (p. 129), not pure apperception itself.

[90] This part of her account is based on a previous article (Freitag/Kraus 2020).

content is generated only when I reflect on the common mark of the apperceived representations, that is, when I attach to them the 'I think': «Attaching the "I think" to an apperceived representation can now be understood as the additional act of conceiving of the original-synthetic unity as a unity, namely as what Kant calls the analytical unity of apperception» (Kraus 2020, p. 109). Kraus's argument goes briefly as follows. The content of 'I think' is minimal because it adds nothing about the subject of the judgement. If I say 'I think the rose is red', usually the mental state that is content of the judgement is not the ascription of a thought to a subject (namely myself): it is rather the objective claim that 'the rose is red'. The same is not the case, though, when I use another pronoun (or a name), e.g. 'she thinks the rose is red': this time I am actually making a judgement about someone else's act of thinking. In the first case, if I omit 'I think' the judgement maintains the same mental-state content. The 'I think' is only added in order to explicitly *express* «the subject realizing the original-synthetic unity that is necessarily required for a significant representation, without determining anything about the subject» (p. 123).

The fact that the transcendental faculty for apperception is the form of consciousness, coherently with her hylomorphic account of mental faculties, means for Kraus that it would be misleading to take it as an act of consciousness itself. However, if on the one hand this avoids the inconvenience of explaining what it can possibly mean for an 'act' not to be empirical/psychological but rather 'transcendental', on the other hand it contrasts with Kant's description of pure apperception as a (surely very peculiar) kind of «consciousness of oneself» (*Anth*. 7: 134n). This means that pure apperception is itself an *act*, and not merely the *form* of an act. Instead of construing pure apperception as a mere form and and denying to it the proper status of an 'act', we should more easily distinguish between 1) the *principle(s)* of apperception, as the *formal* structure or rule under which representations can become part of a cognition of mine; and 2) pure

apperception as an act through which I am aware of my existence*, i.e. of my instantiating a logical self[91].

I believe that characterizing pure apperception as an actual act is compatible with its aprioriness. Indeed, its aprioriness, as I suggested earlier, lies first of all in the very fact that the possibility of such act is a necessary condition of experience. While as an act it is surely in some sense 'psychological' like any other activity performed by the mind, apperceptive self-awareness is not an empirical-psychological condition of thinking, but rather a transcendental-logical condition of objectivity itself. Whereas in empirical apperception a manifold is given to me in inner sense which I then attempt to cognize throughout time as my empirical self and its inner states, the content of pure apperception, the logical I, is not an object of experience and indeed not a individual object at all — not even purportedly, as e.g. the idea of the soul. The logical I is not obtained by means of a synthesis of a manifold of empirical representations standing under the concept <I>, but it is instead co-original with synthetic activity in general. Also, while transcendental apperception *is not* itself the mere form of an act, it *does have* formal features, for the 'I think' entails those conditions for self-ascription of which Kant implicitly provides a «logical exposition» in the Paralogisms chapter — as we have seen in the previous section. In sum, it is possible to avoid conflating pure and empirical apperception — and to leave indeterminate *what* I am as an individual, in favor of the fact «**that** I am» (B157) — without accepting Kraus's reading that pure apperception is the mere form of self-consciousness, maintaining instead that it is itself a peculiar kind of «consciousness of oneself».

I shall now assess another line of interpretation, which contends precisely what my reading of pure apperception takes trouble to deny. As noted above, empirical self-knowledge gives rise to its own set of complex issues.

---

[91] Kraus acknowledges and discusses acts of self-consciousness whose content is neither the empirical nor the noumenal self (supposing that there is such a thing). For example, she reckons a «*Transcendental self-consciousness*», which «is the consciousness of the general form of reflexivity»; and a «*Logical self-consciousness*», namely «the consciousness of oneself as the subject of predication or as the *logical subject*» (Kraus 2020, p. 129). Nevertheless, while these are in some way connected to pure apperception, none of these is itself *equivalent* to pure apperception, for the latter, if I understand her correctly, is nothing more than the *form* of an act of consciousness.

Nevertheless, it is relatively uncontroversial that empirical apperception is supposed to be an access to my self as I appear to myself, and thus only as a phenomenon (cf. B155). Kant then opposes in principle this empirical determination of ourselves as we appear to some putative determination of my self as it is in-itself, while also claiming that the latter is in fact not possible for us since it would require intellectual intuition (cf. B159 and B68). Nevertheless, given Kant's insistence on the difference between empirical awareness of ourselves as phenomena on the one hand, and pure apperception on the other, one way to take such distinction seriously would be to interpret pure apperception as some kind of consciousness of the existence of our *noumenal* self — of our 'self in-itself', as it were — albeit an *indeterminate* type of awareness. It becomes then possible to read the often quoted passage from §25 along this line: «I am conscious of myself not [1] as I appear to myself, nor [2] **as** I am in myself, but only [3] **that** I am» (B157). Consciousness [1] clearly amounts to empirical self-awareness; [2] is supposed to be the determination of my self as noumenon through intellectual intuition, which however is impossible for beings like us. Finally, [3] is what we are aware of in pure apperception, which, according to the interpretative line I am assessing here, should be taken as an indeterminate awareness of my noumenal self.

My aim now is to dismiss the latter interpretative option. In order to do that, let us consider two examples of 'noumenal' reading of pure apperception, namely those advanced by Wuerth and Kohl — which indeed explicitly propose an interpretation of B157 of the kind I have just outlined (see Kohl 2020, p. 397 and 2023, p. 183; and Wuerth 2014, pp. 121-2). As I mentioned earlier, in his book Wuerth (2014) makes extensive use of lecture transcripts and *Reflexionen* to contend that along the late pre-critical as well as the critical period Kant's views on the self are mostly coherent over time — despite some important differences. Very briefly put, Wuerth takes Kant in the *Critical* period to hold the view that through pure apperception I have access to my self as a noumenon. Moreover, while he concedes that in the Paralogisms chapter Kant argues against the possibility to employ schematized categories to cognize the thinking self, on Wuerth's reconstruction Kant would nonetheless defend the legitimacy of applying unschematized categories to our noumenal soul. From the fact that only the

former, unwarranted use of the schematized categories would yield rationalist implications about the immortality of the soul, Wuerth concludes that Kant's criticism of it does not affect the application of unschematized categories to the soul. Wuerth insists that, while the unwarranted use purports to determinately cognize the soul, the second, justified use leaves its nature indeterminate. His proposal faces major difficulties[92]. First of all, contrary to what Wuerth claims (cf. pp. 121-2), nowhere in the Deduction nor in the Paralogisms does Kant explicitly state that pure apperception is the awareness of my self as a thing in itself. While passages such as the previously quoted one at B157 («I am conscious of myself not as I appear to myself, nor **as** I am in myself, but only **that** I am») may be read as *implicitly* making this claim, they can hardly be counted as independent, incontestable textual evidence for it. Secondly, both the negative and the positive findings of the Paralogisms chapter — the former concerning the mistakes of rational psychology, the latter regarding the conditions of self-ascription entailed in the 'I think' — exclude any positive a priori determination of my self as an individual object, and therefore as a noumenal soul.

Wuerth's repeated insistence in using the word 'indeterminate' to characterize our self-consciousness is ultimately not of much help. For instance, even if he excludes that the unschematized category of substance entails the permanence of the object to which it is applied, it would still allegedly establish that such object is the ground of certain accidents. How would this not be a purported *determination* of a thing in itself, namely of my noumenal self as that thing which grounds my inner states and has certain powers? How does this not amount to «a metaphysical determination of the object» (B409), of the type Kant takes to be unwarranted in the Paralogisms? Besides, I have already discussed in the previous chapter how the lack of sensible intuition compromises our ability to make sense of the idea of an individual object (such as, allegedly, my noumenal soul) to fall under a concept. Finally, pure apperception is *immediate*; yet if it were an immediate access to an individual thing in itself, it would come dangerously close to Kant's notion of an intellectual intuition, since we know this immediate

[92] For a more thorough critique of Wuerth, see Proops (2021, pp. 114-22 and section 5.3).

access to myself in pure apperception does not take place in inner sense — contrary to what the pre-critical Kant believed (see e.g. *ML₁* 28: 265).

Kohl (2020) rightly opposes Wuerth's quasi-intuitional model, and yet his proposal turns out equally implausible, at least in my view. According to Kohl, even though, as Kant repeatedly stresses, I cannot achieve any sort of cognition in the «narrow» sense of myself as a determining subject, I can still obtain through apperception a «purely conceptual cognition(W)» — i.e. in the «wide» sense — «of myself as a non-spatiotemporal, noumenal object»[93]. Kohl of course insists that pure apperception does not yet amount to what he calls «*Theoretical Cognitive Determination*» (Kohl 2020, p. 401). He takes this to ensure that there is no inconsistency between the purely conceptual, and therefore allegedly *indeterminate* cognition(W) of my noumenal self at work in pure apperception[94] on the one hand, and Kant's denial that I can determine my self as a spontaneous thinker on the other. However, we have just seen that appealing to indetermination sounds quite arbitrary and unwarranted in cases like this. To be sure, unlike Wuerth, Kohl denies that we have any intuition of my noumenal self — intuition which of course would be required for object determination. Once again, though, I suggest that establishing the existence of a noumenon which grounds my inner states and has certain powers *does in fact amount* to some (however minimal) determination of an object, a determination that in the theoretical domain is unwarranted in absence of some corresponding intuition.

On top of that, it is not entirely clear to me how exactly Kohl characterizes the cognition(W) supposedly yielded by pure apperception. Kohl attempts to

---

[93] «When Kant denies that pure self-consciousness represents the thinking self as an object (e.g., at B407), he uses this narrow notion of an object that designates a determinate object of cognition(N). But if one abstracts from all sensible data for cognition(N), one can still represent an indeterminate "something" via the pure concept of "an object in general" that is the highest concept in all cognition (A290/B346; vol. 28, 544). Accordingly, pure apperception indeterminately represents the thinking self as a "something in general" (A400; cf. A345–46/B403–4; A383; B421). Because pure apperception intentionally relates to a nonsensible ("merely intelligible"; A546/B574) something that is conceived as distinct from my occurrent mental states, it yields cognition(W) of myself as a nonspatiotemporal, noumenal object» (Kohl, 2020, p. 397). See also Kohl (2023), in particular section VI.3.

[94] «[A]scription of spontaneity to the thinking self in apperceptive self-cognition(W) does not suffice to (cognitively) "determine its existence" as a noumenon» (Kohl 2020, p. 410).

establish the «doxastic status» of cognition(W) by proposing what he calls an «inferential reading»:

> Assume that I have cognitive achievements that amount to knowledge. Now assume that I can know that these cognitive achievements presuppose the spontaneous thinking activity of a nonsensible intelligence. If so, I can knowingly infer that "I exist as an intelligence that is … conscious of its faculty for combination" (B158) (Kohl 2020, p. 408).

The question is whether Kohl takes such inferential process to be descriptive of the kind of awareness at work in pure apperception. The just quoted passage as well as Kohl's initial characterization of apperceptive self-consciousness as amounting to cognition(W) lead me to believe that he is indeed suggesting a reading of pure apperception in terms of an act through which I «knowingly infer» my existence as a noumenon. If that is what Kohl has in mind, his reading faces a further objection. As already noted, Kant's instance on the equivalence of 'I think' and 'I exist*', and his (rather enigmatic) characterization of the awareness of my existence* in terms of a «indeterminate perception» or «feeling» strongly suggests that pure apperception itself is by no means an inferential act. In sum, then, Kohl's reading of pure apperception as amounting to cognition(W) of my noumenal self does not seem tenable[95].

More generally, I see at least three considerations that may possibly be advanced in support of some form of 'noumenal' reading of pure apperception. (1) Usually in connection with moral arguments, Kant speaks of a double character of the human being, who belongs both to the empirical and the intelligible world. Regardless of how we understand this complex doctrine, an immediate problem for us here is that Kant appears at times to grant an access to ourselves as inhabitants of the intelligible world precisely through pure apperception. For instance, in the «Clarification of the

---

[95] The suggestion that Kant might be somehow committed after all to some kind of inferential knowledge — perhaps even supported by pure apperception — of my self as a spontaneous being, while deeply problematic, has nonetheless some appeal and would deserve further consideration. However, what I maintain here is that this is not tenable *as a description of what pure apperception amounts to*, as Kohl seems instead to believe.

cosmological idea of a freedom in combination with the universal natural necessity», we read that the human being

> knows [*erkennt*] himself also through pure apperception [*durch bloße Apperzeption*], and indeed in actions and inner determinations which cannot be accounted at all among impressions of sense; he obviously is in one part phenomenon, but in another part, namely in regard to certain faculties, he is a merely intelligible object […]. We call these faculties understanding and reason (A546-7/B574-5)[96].

At a closer inspection, though, the passage is not so straightforward. First of all, what Guyer and Wood (but also Kemp Smith) translate as «pure apperception», in the original German is «*bloße Apperzeption*», which could be more properly translated as 'mere' apperception[97]. That said, it is nonetheless fair to note that when Kant mentions apperception without further specification he usually has the pure one in mind. Secondly, though, it is not clear from the phrasing whether what I «*erkenne*» «*durch bloße Apperzeption*» is myself as an intelligible object, or rather only my spontaneous faculties. A final layer of complexity is added by the fact that the quoted passage belongs to a part of the *Critique* that is concerned with freedom, which however Kant did not manage to revise in the second edition of the *Critique*[98]. This is relevant for the following reason.

In the third section of the *Groundwork*, we find a similar argument from the spontaneity of certain human faculties to the twofold nature of human beings. There Kant writes:

> with regard to what there may be of pure activity in him (what reaches consciousness immediately and not through affection of the senses) he

---

[96] This passage is indeed cited by Kohl (2020, pp. 395-6).

[97] Indeed, the Italian translations of Giorgio Colli, Pietro Chiodi and Constantino Esposito translate 'bloße' with '*semplice*', i.e. simple in the sense of 'mere'. 'Mere' is also the term chosen e.g. by Jauernig when quoting this passage. She nonetheless takes this passage to be one of those cases in which Kant «sometimes even sounds as if he believes that in self-consciously thinking we are aware that we have a supersensible part that belongs to the transcendental level of reality» (Jauernig 2021, p. 280).

[98] As Kant himself acknowledges (BXXXVIII), for reasons of time the changes in the B-edition do not go beyond the Paralogisms.

> must count himself as belonging to the *intellectual world,* of which however he has no further cognizance (4: 451).

One may of course argue that the immediate consciousness alluded here, that allows us to think ourselves as belonging to an intelligible world, is indeed pure apperception, and yet this is far from explicit. More importantly, interpreters usually agree that Kant's view on our access to freedom changes to some relevant degree by the time of the *Critique of Practical Reason*, where freedom's *ratio cognoscendi* becomes the famous 'fact of reason', i.e. our awareness of the moral law[99]. We may then wonder to what extent the passage at A546-7/B574-5 still represents Kant's considered view at the time of the B-edition of the first *Critique*. In any case, while it cannot be denied that these passages are tricky to account for on the view I am defending here, it is sufficient to note that they do not make a crystal clear case for a 'noumenal reading' of apperception either. Given that the Deduction and the Paralogisms do not present explicit textual evidence that in apperception we are aware of our noumenal self — the Paralogisms go indeed quite in the opposite direction — it seems to me that priority should be given to Kant's more systematic claims, rather than to those scattered pieces of an entirely undeveloped and hardly consistent view.

(2) Another consideration that can be made is that, at the end of the day, it cannot be denied that Kant's moral philosophy ultimately leads him to an explicit commitment to the existence of the self as a noumenal being endowed with freedom and, eventually, immortality. While this could be taken as a decisive clue about Kant's final view on the ontological status of the self, his moral arguments in the second *Critique* do not per se concern his account of pure apperception. The fact that they rely on the logical possibility, disclosed by transcendental idealism, of thinking a non-spatiotemporal layer of reality is harmless for us here.

(3) Finally, it may be argued that Kant's account of faculties in general, and in particular of the spontaneity of thought entails a commitment to the existence of a noumenal being, however it may be like, that grounds these mental powers. Unfortunately, Kant does not provide us with an explicit and

---

[99] For a reconstruction of this change in Kant's view, see e.g. Ameriks (2000, ch. 6) and Timmermann (2010).

definitive view on these crucial issues. My only response for now is, once again, that considerations on the non-empirical ground of the activity need not affect directly the Deduction and pure apperception, for according to the type of interpretation I am suggesting here they belong to another level of discussion. Now, Kant writes in the WRP that the «logical self», which is an «*a priori* representation» (20: 270), «does indeed point to the subject as it is in itself» (20: 270-1): indulging in speculation, we may read this claim as to suggest a connection and yet at the same time a *distinction* between the (transcendental-)logical standpoint of and the Deduction, where the logical self plays its role, and the more metaphysical quest for the non-sensible ground of our spontaneous activity.

***

The main goal of the chapter was to further support my deflationist account of Kant's view of things in themselves by considering the special case of the self in pure apperception. I started with an overview on the principle(s) of apperception in §16 of the Deduction. I then presented my solution to the question of what pure apperception is about. On the positive side, it is the awareness of the fact that I bear the status of an 'I', namely the identical subject to which I must be able to ascribe my thoughts if they are to figure in cognition. The 'I' of the 'I think' is therefore not a concrete, individual object, but instead a status that I must be able to attribute to myself if cognition is to obtain. I then maintained that pure apperception is not a mere form, but rather an act, whose possibility is a normative demand if experience is to obtain. On the more negative side, my main claim is that pure apperception must not be intended to amount to an awareness of my self as a noumenon. I reinforced my account through a quick confrontation with a couple of interpretations where these markedly differ from mine. Finally, I assessed three appealing considerations in support of some type of 'noumenal' reading of pure apperception and eventually of the self. I nonetheless concluded that these are in fact not decisive at all, especially as far the content of pure apperception is at issue.

4

The Specter of Monadology: Kant's Reappraisal of the Concepts of Simple and Substantiale

A series of passages in Kant's works raise a question of no secondary importance: Is Kant committed to some sort of *noumenal monadology*? This question, which is more pressing than one imagines, will be the guiding thread of the present chapter. By submitting a negative answer to this question, I aim to provide further evidence against some of the 'metaphysical' trends in Kantian scholarship. More than that, through a clarification of Kant's treatment of the notions of <*simple being*> and <*substantiale*>, I hope to bring to the fore new elements for a deeper understanding of Kant's stance towards noumena, and in particular of Kant's complex efforts towards a critical reappraisal of some of the tenets of his predecessors' doctrines.

Section 4.1 submits an argument that is functional to the remainder of the chapter, for I argue that Kant is not theoretically committed to the existence of the unconditioned among things in themselves. Section 4.2 presents an overview of Kant's unsettling claims about monadology, in particular about simplicity. In section 4.3, I focus on an illuminating passage from *On a Discovery* to illustrate Kant's construal of <simple being> as the negative concept of an unconditioned condition of the composite. Section 4.4 assesses the claim that monadology is analytically true of things in themselves. Some version of this view is seemingly held by Kant. I engage with Oberst's view on the topic inasmuch as it is representative of a 'metaphysical' line of interpretation. I then oppose to it an alternative, deflationist construal of Kant's commitment to the analytic truth of monadology. Section 4.5 revolves around the much overlooked notion of <substantiale>. I begin by showing how the substantiale exceeds the limits of our faculties. I subsequently focus on a passage from the *Prolegomena* in order to argue that this notion plays an important role parallel to that of <simple being>. I thus submit another analytic proposition about noumena, this time involving the concept of <substantiale>. After considering some complications arising from my account, section 4.6 briefly discusses how the notions of <simple being> and

<substantiale> converge in the concept of a monadic soul. The chapter ends with a suggestion about the way in which the latter concept, and in particular the notion of <substantiale> as I characterized it, may help us to rethink the noumenality of the soul in deflationist terms.

4.1 Prelude: Reason, PSR and the Unconditioned

Before starting our investigation on the main topic of this chapter, namely Kant's ambiguous attitude towards monadology, let us start with a different, but closely related issue. In the following, I shall argue that Kant is *not* committed to the view to the effect that, at least in the noumenal domain, there exists (or there would exist) something unconditioned for every series of conditions. The view that I am denying here is one that takes Kant to be committed to a version of the Supreme Principle of pure reason (henceforth SP) as descriptively true of things in themselves. Since, on Kant's account, the Supreme Principle itself is not limited to the noumenal domain, but it is also not of *constitutive* use in experience, let us call such restricted, allegedly constitutive version of the principle a 'noumenal' version of the Supreme Principle (call it NSP). The issue is relevant first of all because, if the NSP were true, we would be thereby in possession of a principle valid for things in themselves, which would amount to cognition (at least in a broad sense) of the noumenal domain. This clearly goes against the account of the limits of discursive understanding that I am defending here as an interpretation of Kant. Secondly, establishing that the NSP is unwarranted is functional to an argument that I shall present below in this chapter (section 4.4). Indeed, I will attempt to show there that the proposition stating the existence of *simple beings* and *substantialia* in the noumenal domain is analytically true only on the assumption of certain *unwarranted* premises, one of which is the NSP.

Kant's views on the unconditioned and the Supreme Principle have been object of growing interest in Kantian scholarship at least since Grier's (2001) seminal book on transcendental illusion. In the following, I will not attempt at a systematic account of reason, nor of the exact import of the SP. I will draw on the quite uncontroversial insight that the SP is not constitutively

valid of experience, while also leaving the possibility open that it may have a useful regulative function, a topic that I will set entirely aside.

Pure reason's SP holds a complex connection to the Principle of Sufficient Reason (PSR)[100]. Kant confronts himself many times with the PSR since the *New Elucidation* (1755). In the critical period Kant arrives at a firm denial of its unrestricted validity as a metaphysical principle concerning all possible things[101]. It has often been argued, correctly I believe, that the PSR survives in the vest of the Second Analogy[102], namely Kant's synthetic a priori principle of causality, which is valid exclusively within the spatiotemporal domain, i.e. within experience[103]. Now, the reason why the relation between the PSR and the SP is complex is because the SP is in some sense *based on* the PSR, but at the same time *excluded* by it. Kant's formulation of the SP reads: «when the conditioned is given, then so is the whole series of conditions subordinated one to the other, which is itself unconditioned, also given (i.e., contained in the object and its connection)» (A307-8/B364). The SP involves the PSR inasmuch as it entails the idea that each member of a regressive series of conditions is in turn conditioned, that is, grounded in another member, namely its condition. However, the SP is incompatible with the PSR inasmuch as the PSR excludes that something can be unconditioned, i.e. exist

---

[100] Some accounts of the relation between these two principles can be found for example in Willaschek (2018, section 3.4), Lu-Adler (2021) and Schafer (2023, ch. 5).

[101] See *OaD* 8: 193-8 for a distinction between a logical (formal) and a transcendental (material) version of the PSR that concern either propositions or things respectively. The transcendental, unwarranted version of the PSR is formulated there as follows: «*every thing must have its ground* [*Grund*]» (*OaD* 8: 193). I would like to take issue here with a reading that is attentive to the limits of the PSR, albeit, however, for an incorrect reason. Hogan (2022) argues that Schopenhauer's praise of Kant's Transcendental Aesthetic is guided by an insight that ultimately allows us to appreciate why Kant could hold on to the mere subjectivity of space and time against Trendelenburg's 'neglected alternative objection' (on the latter, see section 5.3 below). To put it succinctly, according to Hogan, on the one hand space and time are representations that involve the PSR inasmuch as their parts determine each other reciprocally; on the other, Kant firmly maintains that things in themselves are excluded from the PSR. Against Hogan, I take it that the restriction of the PSR derives from the ideality of space and time rather than the other way around, for it follows from the conditions of validity of the Second Analogy — which holds for sensible reality alone and can only be valid a priori precisely *because* space and time are forms of intuition.

[102] This is indeed suggested by Kant at A200-1/B245-6. On the connection between the PSR and the Second Analogy, see for example Longuenesse (2005, ch. 7) and Lu-Adler (2021).

[103] Cf. section 1.5 above.

without a ground, whereas the SP affirms precisely the existence of something unconditioned — be it the whole series or a member of it[104]. Note that, given Kant's denial that something can ground its own existence[105], he takes the PSR to be incompatible with the existence of an ungrounded being such as God[106]. Therefore, perhaps because Kant takes the PSR to conflict with reason's interests, he has replaced the PSR with the SP, where the latter is a version of the former that has been «revised in order to accommodate the irreflexibility of grounding and thus to allow for something unconditioned», as Willaschek (2018, p. 101) suggests. What matters for the present discussion is that the SP (and so the NSP) relies on there being a necessary grounding connection among conditioned things.

Now, as I said, several interpreters, for different reasons, take Kant to be committed to what I have called NSP, namely a restricted version of the SP that is supposedly valid of noumena[107]. More precisely, one may find Kant to hold that the NSP is true and (either consequently or not) that things in

---

[104] To be sure, Kant's formulation of the SP is problematic, for it appears to exclude a case that he takes undeniably seriously throughout the Transcendental Dialectic, namely the possibility that the unconditioned in a series be a member of the series, and not just that the series as a whole be unconditioned. Indeed, the notions of soul, simple part, or freedom, for instance, are arguably purported unconditioned conditions. This difficulty has of course been noted by the interpreters, such as, for example, Willaschek (2018, section 3.3) and Schafer (2023, ch. 5). For simplicity, consider Schafer's reformulation of the SP, which includes both scenarios mentioned above: «When some X conditioned in respect R is given, then either (i) the whole series of R-conditions involving X, which is itself unconditioned, is given or (ii) some other (non-R) condition(s) of X is given, which is sufficient to complete reason's unity» (Schafer 2023, p. 156).

[105] «[T]he ground of the existence of a thing, as real ground, must always be distinguished from this thing and this must then necessarily be thought as dependent upon another» (*OaD* 8: 198).

[106] On the incompatibility between the PSR and the unconditioned, and consequently God, Kant writes: «the proposition itself [i.e. the PSR], in the unlimited universality in which it there stands, is, if it is to be valid of things, obviously false; for according to it there would be absolutely nothing unconditioned; but to seek to avoid this embarrassing consequence, by saying of the supreme being that it does, indeed, also have a ground of its existence, but that this lies within it, leads to a contradiction; for the ground of the existence of a thing, as real ground, must always be distinguished from this thing and this must then necessarily be thought as dependent upon another (*OaD* 8: 198).

[107] Some examples are Jauernig (2021, section 6.6), Schafer (2023, ch. 5), Proops (2021, section 10.6), Willaschek (2018) and most thoroughly and radically Stratmann (2022). See also my discussion of Oberst (2018) in the remainder of the present chapter.

themselves are actual. Or, alternatively, one may advance the weaker claim that the NSP is somehow implicitly contained in Kant's conception of noumena without thereby assuming that things in themselves exist; that is, one may hold NSP as objectively valid only on the assumption that noumena exist. Against all that, I will attempt to show that the NSP too is unwarranted even in its weaker version.

I will consider (and reject) two main reasons to ascribe the NSP to Kant. Let us start with (1) a more speculative argument in favor of such ascription. According to such argument, Kant's conception of noumena already necessarily entails that these conform to the understanding taken in the broad sense — call this the Conformity Claim (CC). Consequently, whereas the SP is not constitutive of phenomena, if intelligible entities exist they necessarily conform to reason's (N)SP. Here, however, are the difficulties with this position. It must first of all be appreciated that principles of *reason* and principles of *understanding* are significantly different for Kant, both with respect to their validity (regulative vs. constitutive respectively), as well as to what they are supposed to achieve. While both kinds of principles are required for experience to obtain as a lawful, systematic connection among phenomena, their functioning differs significantly, e.g. in their relation to intuitions. Therefore, simply stating that principles of understanding and reason would indifferently apply to intelligible entities is quite simplistic. Let us see why.

In Chapter 1, I have shown that characterizing non-sensible entities as objects of some intellectual intuition is for Kant our best, and yet still insufficient attempt of making sense of what it would mean for an intelligible being to be an existing individual. Consider how the Conformity Claim behaves when we conceive of noumena as objects of an intuitive intellect. Both principles of reason and of understanding crucially involve concepts, whereas intellectual intuition would make no use of them: What would it even mean, then, for objects of intellectual intuition to be governed by principles involving concepts? One may insist here that our principles would be *our* (a finite rational being's) way to account for what is at issue within an intuitive, i.e. non-conceptual intellect. Now, not only things start getting quite chaotic here, but note also that noumena are non-sensible, i.e. non-spatiotemporal beings, whereas Kant's synthetic a priori principles of the

understanding involve some reference to our sensibility. For instance, the second analogy, or principle of causality bears the title: «**Principle of *temporal* sequence according to law of causality**» (B 232, italics mine). Therefore, affirming that noumena, i.e. non-spatiotemporal entities would eventually conform to *Kantian* synthetic a principles of pure understanding is in fact a non-sense — at least if the principles we have in mind are the synthetic a priori ones presented by Kant in the Analytic of Principles.

Note however that, while it is non-sensical to claim that 'rationalists' take things in themselves to conform to *those synthetic a priori principles of pure understanding that Kant has exposed*, my point is nonetheless consistent with the fact that Kant accuses the 'rationalists' of mistakenly believing that things in themselves — or rather, all possible things 'in general' — conform *to what they take to be* the principles of the understanding, e.g. the Leibnizian principles attacked in the Amphiboly. Indeed, one of Kant's criticism of his 'rationalists' predecessors involves the charge that they mistake logical principles establishing connections among concepts for metaphysical principles about all possible things[108]. As I see it, this error has one of its roots in the fact that, given our unawareness of any criteria for the individuation of non-sensible beings (see Chapter 1), in attempting to gain insight into the latter we find ourselves forced to play with empty *concepts*. Yet unlike the rationalists who believe to cognize non-sensible beings when all they are looking at are, as it were, 'hypostatized concepts', Kant is ready to admit our ignorance of the criteria for numerical distinctness when we depart from sensibility and the consequent failure of metaphysical investigations into the constitution of non-sensible beings.

Having established that Kant does not take noumena to conform to the principles of understanding, we can now see that things do not get much better with principles of reason, such as the SP. Kant maintains that principles of reason are only *regulatively* valid. So first of all, we must detach those principles from the 'validity status', as it were, that Kant ascribes to them. The CC must thus be properly configured here as to mean that a noumenal world would conform to principles of reason in the sense that

---

[108] Besides the Amphiboly, see for example (again) Kant's warning against turning the formal version of the PSR into its material and unrestricted version at *OaD* 8: 193-8.

those principles would be *constitutively* valid of it. Since it can be argued that the regulative status of rational principles available to us is a consequence of the finitude of our human condition, let us call an 'infinite (or purely) rational being' the subject whose objects would supposedly immediately conform to the principles of its own pure rational faculty. Now, it could be argued that our best way of making sense of an intelligible world as a given and interconnected totality is to construe it as the object of an infinite rational being. As it was the case with intellectual intuition, merely considering this logical possibility is a harmless, and perhaps useful exercise. However, we have no theoretical access to such an infinite rational being nor to some intelligible world conforming to it. Hence, it is unwarranted to claim that those are *really* possible, let alone actual. Moreover, I see no grounds for us to exclude once and for all the possibility of some intelligible world that *is not* the object of an infinite rational being whose principles are constitutive of such world. This amounts to say, I take it, that construing a non-sensible world as already conforming to the constitutive principles of an infinite rational being is hardly a conceptual necessity. Therefore, the NSP does not necessarily obtain, not even as an analytic truth about things in themselves[109].

So far so good, but what if it was necessary for us to assume, even on theoretical grounds, the existence of some infinite, purely rational faculty and of its unconditioned objects? This is the thrust of Schafer's (2023) proposal —

[109] Messina (2021) is committed to the CC against a by now familiar background. Starting from the much employed distinction between negative and positive noumena (to be characterized as non-sensible beings, and as objects of pure understanding alone respectively), he finds Kant to be committed to the existence of the former but not of the latter. Only the latter though, i.e. positive noumena, are to be construed as necessarily conforming to principles of understanding and reason: «positive noumena for Kant are beings conceived in a certain way — as fully conforming to the intelligibility constraints of the understanding operating in tandem with principles (such as the Principle of Sufficient Reason) and interests of reason, including its interest in completeness» (Messina 2021, p. 13). Similarly, Willaschek writes that «*if* there are *noumena* in the positive sense, they have to conform to principles of reason, including the Supreme Principle, because by definition they have to be part of a fully rational order of things» (Willaschek 2018, p. 154). My account differs in that I take conformity between noumena and reason, with the caveats exposed above, as no more than a logically possible way for us to make sense of a world of non-sensible entities. A positive *use* of <noumenon> amounts then to the unwarranted move of taking individual noumena to be actual or really possible, or of assuming that an intelligible world exists as composed of lawfully interconnected noumena (on the advantages of my account of the two uses of 'noumenon', see section 1.4).

which involves a teleological reading of mental faculties. Here is Schafer's book-long argument reduced to its very core. An intelligible being (or what he and other would characterize in terms of 'positive noumenon') would be the object of a purely rational being whose infinite faculty of reason would be the 'properly functioning' version of human finite, defective type of reason. For such purely rational beings, rational principles would turn out to be constitutive. Human faculties, including reason, intrinsically strive for their full realization, i.e. to function 'properly'. Such striving is only legitimate on the assumption that a proper functioning is at least possible. Analogously, our search for the unconditioned is only legitimate as long as the unconditioned itself is possible. Now, Kant is adamant that we can achieve no proper cognition of unconditioned beings. Schafer's solution is then to argue that the unconditioned is for us object of 'theoretical belief', i.e. a mode of 'assent' (or *Fürwahrhalten*, as Kant calls it) justified on subjective, but not objective grounds[110]. Schafer maintains that the NSP would accordingly be valid on merely subjective grounds.

Clearly, there would be much to discuss here, and it is impossible for me to attempt at assessing Schafer's suggestive proposal. In order to support my view that the NSP is unwarranted, consider just that, first of all, there is no explicit textual evidence of Kant being committed to the unconditioned in the form of theoretical belief — a notion that in fact has little or no explicit role outside the Canon of Pure Reason[111]. At any rate, though, even if we accepted Schafer's suggestion, the decisive question for us here would be whether theoretical belief in some x requires us to have established x's real possibility, or just its logical possibility instead. While I have not developed a detailed view on such issues, it seems to me that the requirement of x's real possibility is unnecessarily overdemanding. I think that, for us to hold a warranted belief in x (in the Kantian sense), it is enough for us to establish that x is

[110] Cf. Section Three of the Canon of Pure Reason. For an influential discussions of theoretical belief in Kant, see Chignell (2007) and Stang (2016, ch. 9).

[111] By this, though, I do not mean to simply deny any legitimacy to such plausible interpretative solution.

logically possible, *and* that its *real impossibility* is theoretically *indemonstrable*[112]. After all, a switch in modality, as it were, is precisely what is at issue in the very act of belief: believing in a possible x, means in fact taking x to be actual (or really possible[113]) on subjective grounds alone. As a result, I maintain that we cannot take Kant to hold the NSP — at the very least not on objective theoretical grounds.

(2) One may advance a second reason to hold the NSP, that is, it could be argued that the NSP is entailed in Kant's characterization of transcendental realism. Let us start from the less controversial fact that a distinctive way in which Kant construes the errors of his transcendental realist predecessors in the Transcendental Dialectic, in particular in the Antinomies, is by claiming that those philosophers mistake appearances for things in themselves[114]. This fallacious ontological position gives thus rise to the fatal mistake of assuming the unconditioned to be actually given. The question now is: What does that mean? On a construal of this story, Kant would assume that things in themselves are governed by the NSP. The transcendental realist's mistake would consequently be that of taking a principle valid for things in themselves alone (i.e. the NSP) as if it were also valid for what is given to us in experience, thus generating the equally misleading assumption that the

---

[112] Although I cannot provide an extensive argument for it, I hold this to be the thrust of Kant's strategy in the Transcendental Dialectic with regard to those concepts that will be characterized in the moral domain as objects of practical cognition or belief. Hence I distance myself from Proops' claim that, in the Third Antinomy, Kant does not even take to have established the logical possibility of freedom as «that concept might rather contain a hidden internal contradiction» (Proops 2021, p. 279). See Kohl (2023, p. 26 fn. 36) for a criticism of Proops on this point. Although I disagree with Proops's claim from an exegetical standpoint, the question could be raised of whether establishing the logical possibility of a concept is in fact more problematic than Kant suspects.

[113] For example, I may believe in the *actuality* of Kant's ghost haunting my house, but also in the *real possibility* of not-yet-actual events such as miracles, my future reincarnation, or an imminent solution to the climate crisis.

[114] See, for example, A491/B519 and A740/B768.

unconditioned is in fact given to us. In sum, on this picture Kant would implicitly endorse the NSP[115].

However, a more plausible reconstruction of what taking appearances to be things in themselves amounts to is available — a reconstruction that rather supports my rejection of the NSP. Such conflation amounts to take the objects which are given to us in experience, *along with the principles governing them*, as existing in themselves[116]. Consider the following passage to be found in a paragraph of the *Prolegomena* where Kant is explicitly discussing the Antinomies:

> If (as normally happens) we think of the appearances of the sensible world as things in themselves, *if we take the principles of their connection to be principles that are universally valid for things in themselves and not merely for experience* (as is just as common, nay, is unavoidable without our *Critique*): then an unexpected conflict comes to light, which can never be settled in the usual dogmatic manner, since both thesis and antithesis can be established through equally evident, clear, and incontestable proofs — for I will vouch for the correctness of all these proofs – and therefore reason is seen to be divided against itself, a situation that

---

[115] Some version of this implicit commitment seems suggested by A416/B444 (quoted e.g. by Stratmann 2022 and Schafer 2023, p. 166) and A498/B526 (quoted e.g. by Proops 2021, pp. 258-9), passages which may be taken to support the claim that the (N)SP is descriptively valid. However, a more deflationist reading of these passages is also available, namely one that concedes only that the NSP is implied by the transcendental realist insofar as he construes things in themselves (and thus things in general) as accessible through (and as to conform somehow to) pure understanding alone. Such a reading englobes some useful insights on (what I called) the Conformity Claim from authors as for instance Willaschek (2018) and Jauernig (2021). However, at this point a proper discussion of the case at issue would be excessively convoluted. For example, depending on how we characterize what such transcendental realist takes to be the principles governing understanding, the resulting consequences for noumena may turn out — in light of Kant's doctrines — either non-sensical, metaphysically sterile, or merely logically possible. To settle the issue, it suffices here to say that these passages are ultimately harmless for my aims. For they need not be read as to entail any commitment on Kant's part to the truth of the NSP as a description of existing things in themselves, nor as a feature that must necessary be included from the start in the concept of a non-sensible being. That said, I believe things go more smoothly when we focus — as I shall do in a moment — on an alternative and more plausible reconstruction of the transcendental realist's mistake.

[116] This construal fits quite naturally Kant's doctrine of the impossibility to extend cognition beyond possible experience. To be clear my primary aim here is not to provide a novel interpretation of the transcendental realist mistake in general, but rather to defend my denial of the NSP against possible objections.

> makes the skeptic rejoice, but must make the critical philosopher pensive and uneasy (*Prol*. 4: 339-40, italics in the sentence mine).

Here Kant tells us explicitly that taking the *principles of the connection* among appearances — which plausibly mean the synthetic a priori principles of the understanding — to be universally valid for things in themselves is a crucial part of the mistake that leads reason to its internal conflicts, e.g. the antinomies. Reason's conflict, we know, involves the transcendental realist assumption that the unconditioned must exist. Hence my point is that Kant's attack on the transcendental realist erroneous commitment to the SP leading to the antinomies does not rely on the assumption that the (N)SP is valid of things in themselves. Instead, Kant's diagnosis of the mistake committed by transcendental realism is that the latter illicitly extends to all things in general those principles that are a priori valid only for appearances. Note that among the principles that govern appearances is of course the Second Analogy, which, as suggested earlier, is a legitimate version the PSR. Therefore, what happens is that transcendental realists endorse an unrestricted version of the PSR, one which would eventually be valid for all existing things regardless of whether these are intuited by us or not. If the PSR holds unrestrictedly, then «*every thing must have its ground* [*Grund*]» (*OaD* 8: 193), that is, for every conditioned thing there exists its condition.

At this point it may be objected that this is not yet equivalent to the (N)SP, for we have seen Kant that takes the PSR to exclude in fact the existence of any unconditioned condition. However, things must be seen here not from the perceptive of Kant's own criticism of the PSR, but rather from that of the very position criticized by Kant (e.g. in the *OaD*), a position that is *not able* to

detect a decisive incompatibility between the PSR and the existence of unconditioned conditions[117].

In conclusion, there is good reason to *reject* those construals of both the Conformity Claim and of Kant's diagnosis of transcendental realism's mistake which would allegedly imply the truth of the NSP. We are therefore not warranted in maintaining the existence of some unconditioned noumenon, not even as a fact necessarily included in the concept of noumenon. The existence of the unconditioned is rather an assumption of the transcendental realist, who takes the principles connecting appearances to be unrestrictedly valid. The negative result of this section will turn out especially important for our discussion on the analytic truth of monadology (see section 4.4). Before that, let us introduce the problem of Kant's ambiguous attitude towards monadology.

4.2 Kant's Ambiguous Statements on Monadology and Simplicity

Kant's vocabulary in the *Lectures on Metaphysics* and elsewhere make it often seem as he believed that we know that (some or all) things in themselves are *simple*. More specifically, Kant seems to entertain the view that Leibniz's *monadology* is somehow correct about things in themselves. To be precise, these issues concern also passages in his published texts, some of which will be examined in the following. Nevertheless, the lectures are significantly more puzzling in this regard. First of all, though, it must be

[117] The construal of the transcendental realist mistake that I attempted to reject plays instead a relevant role in Stratmann's proposal: «insofar as the dogmatic metaphysician is susceptible to taking appearances for things in themselves, she is susceptible to illegitimately supposing that the supreme principle has existential import for (what are in fact) appearances, and thus to illegitimately inferring the existence of something unconditioned among appearances» (Stratmann 2022, p. 529). Especially crucial for Stratmann's complex argument is the claim that, for Kant, «the assumption that things in themselves would have to be completely determinate underlies the Antinomies chapter's» indirect argument for idealism (p. 530). This is important because Stratmann needs things in themselves to be so characterized in order for the (N)SP to analytically obtain of thing in themselves. However, unlike Stratmann, I take it that the acceptance of the Principle of Thoroughgoing Determination as descriptively valid for things in themselves (or, as it is equivalent for the 'rationalist': for all things in general conceived through pure understanding alone) is a mistaken assumption which Kant attributes to the dogmatic metaphysician, rather than an assumption made by Kant himself.

reminded that, despite their enormous and indisputable value, the transcriptions of Kant's lectures are more problematic than Kant's published writings for several reasons[118]. One is that Kant conceived all his lectures as a sort of commentary and discussion of a textbook. Kant's *Lectures on Metaphysics* have Baumgarten's *Metaphysica* as a continuous counterpart. As a consequence, in many passages it is not clear whether Kant is expressing his own views, or rather discussing Baumgarten's position. Indeed, the notes mention only seldom (and often suddenly) Baumgarten, or 'the author'. In addition to that, it is not unusual to find contradictory statements in the range of a few pages, and it is once again hard to establish whose fault is it. Having said that, in the following discussion I will avoid appealing to those intrinsic difficulties of reading Kant's lectures. I shall now first illustrate the problem with some examples; I will then suggest solutions which in turn have an extremely relevant meaning for our understanding of Kant's stance towards noumena.

In the *Metaphysik Mrongovius*, Kant is reported to have said that we can hold a substantial composite to consist of simple substances (monads), «*insofar as it is a noumenon*» (*MM* 29: 827-8; for similar claims, cf. *MM* 29: 859 and *MD* 28: 663-4). He says also that it is correct to maintain of noumena that the simple remains if «I set aside all relation» (29: 828). Things get even more complicated as Kant is reported to have said, at the beginning of the section on cosmology, that we obtain the notion of simple elements by considering the «world of the understanding» (29: 850); he then specifies that we can cognize of that world no more than what is included in its definition (cf. 29: 850). Nevertheless, he makes it clear that the latter claim does not imply that such concept of world is a mere fantasy; instead, it is «this world as I think it through the understanding». A similar line of reasoning can be found in the *Metaphysical Foundations of Natural Science*, where we read that monadology is

> an intrinsically correct *platonic* concept of the world [*an sich richtiger platonischer Begriff von der Welt*] devised by *Leibniz*, insofar as it is considered […] as a thing in itself, and is merely an object of the understanding, which, however, does indeed underlie the appearances of the senses. Now the composite of things in themselves must certainly

[118] On these issues, see for example Clewis (2015).

> consist of the simple, for the parts must here be given prior to all composition. But the composite in the appearance does not consist of the simple (4: 507-8).

These passages might be taken to support the reading of those interpreters who are willing to allow for some kind of positive claim concerning the nature of things in themselves. For instance, Langton (1998, p. 202) quotes the *MFNS* passage as a major evidence for ascribing to Kant the view that the simplicity of noumena is a real state of affairs, and thus to find him committed to some sort of noumenal monadology[119]. And Ameriks (1992) lists the aforementioned passages from the *Lectures on Metaphysics* in support of his claim that Kant is committed to simple beings outside the spatiotemporal domain[120]. I shall now show that this is not the case. At least two sets of issues arise from the previously quoted passages: (I) Is Kant positively determining and ascribing properties to things in themselves? Is Kant claiming that existing things in themselves fall under the concept of <simple being>? If the answer is no, where does the concept of <simple being> come from? (II) What does it mean that monadology provides an intrinsically correct concept of the world? What does it mean that all we are able to know about the world of understanding is entailed in its definition?

I shall start by addressing the first problem, which is also the most relevant. Right below the first passage quoted above from the *Mrongovius* lectures on metaphysics, we read that, in order to «speak of simple beings we must go beyond the world of the senses, *but then we have no proof for the objective reality of our concept, for we can give no example*» (*MM* 29: 828, italics mine). Significantly, these words echo a line of reasoning that can be found in

---

[119] In Chapter 2 I mostly focused on the similarities between Langton and Allais. One of the many differences between their accounts emerges, however, in their views about monadology, for Allais (2015, p. 73) explicitly denies that Kant is committed to the simplicity of things in themselves.

[120] «The departure from traditional ontology comes not in a denial of simple beings but in a refusal to allow them to be understood as directly spatiotemporal or as such that spatiotemporal properties can be considered as in principle derivable from the concept of those beings. Given the conclusion of the First Antinomy that the spatiotemporal domain is merely phenomenal, this means not that simple beings are to be dismissed ontologically but rather that they are saved — even if their individual determination is ruled out for us» (Ameriks 1992, p. 261).

a previously quoted passage from the Amphiboly (see section 2.3). There Kant warns us that «no example can be derived from anywhere except experience», and that therefore we cannot understand whether «*realitates noumena*» represent «something or nothing at all» (A282-3n/B338-9n). This is the case of the concept of a reality in which no real opposition is possible, discussed in the Amphiboly, but also of the concept that concerns us here, namely that of a simple being. Therefore, Kant's claim that in the latter case «we have no proof for the objective reality of our concept, for we can give no example» is a strong enough statement to legitimize some serious suspicion, to say the least, towards the apparent correctness of monadology. Indeed, in the A-Paralogisms, Kant writes: «even the fundamental concept of a **simple nature** is of such a kind as cannot be encountered anywhere in experience, *and hence there is thus no path at all by which to reach it as an objectively valid concept*» (A361, italics mine).

Another reason to be cautious concerns Kant's choice, in the aforementioned *MFNS* passage, to label Leibniz's monadology a «*platonic* concept of the world» (4: 507-8). It is certainly true that Kant aims through this expression to distinguish Leibniz's doctrine from the Wolffian 'physical' monadology. Moreover, there is no denying that Kant is especially critical towards the latter, for he now considers the notion of physical monad to be inappropriate compared to Leibniz's original usage of the term[121]. This need not be read, however, as a commitment to Leibniz's monadology on Kant's part. In discussing the topic in light of the relevant historical background, Ottaviani (2018) stresses — in my view correctly — that the label '*platonic*' that Kant ascribes to Leibniz's monadology is hardly indicative of approval. For besides his famous praise of Plato's doctrine of the transcendence of ideas[122], Kant is overall very critical towards Plato's «mystical system», which «demanded that the true objects be merely intelligible, and asserted an intuition through pure understanding not accompanied by any senses» (A854/B882).

---

[121] See for example the Remark to the antithesis of the Second Antinomy, and the essay *On a Discovery*, which I shall examine below.

[122] At the beginning of the Dialectic (A313-9/B370-5) and elsewhere (e.g. A 568-9/B596-7).

4.3 The Simple as Unconditioned

In the previous section I attempted to show that the question raised with issue (I) should be answered negatively: Kant is *not* stating that (some or all) things in themselves fall under the concept <simple>. But if that is not what he is doing, what is then more exactly the concept of a simple being and how does Kant obtain it? In order to solve this puzzle, it might be helpful to take into consideration some remarkable pages from Kant's essay *On a Discovery* (8: 208-11)[123]. In the passage I shall discuss, Kant accuses Eberhard — who is the main target of the essay — of having misinterpreted his distinction between appearances and things in themselves. He then continues by reminding that the simple does not belong to the corporeal world, and that our impossibility to experience it does not depend on the mere degree of sharpness of our sensibility. Then he writes: «if reason thinks a composite of substances as thing-in-itself (without relating it to the special character of our senses), it must absolutely conceive it as composed of simple substances». At a first glance, this statement would seem to be a further evidence for holding things in themselves to be positively determined as composed out of simple entities. However, Kant's view seems again more complex than that: «no recourse remains but to admit that bodies are not things-in-themselves at all, and that their sensory representation, which we denominate corporeal things, is nothing but the appearance of something, which as a thing-in-itself

[123] For more on the *OaD*, besides the aforementioned Ottaviani (2018), see also Allison (1973), La Rocca (1994) and Betti (2024). Allison's «Historical-Critical Introduction» to the volume dedicated to the *Kant-Eberhard Controversy* is one of the most relevant works on the *OaD*. There he occasionally engages with Kant's stance towards noumena in a way that is consistent with his later work on Kant's transcendental idealism (i.e. Allison, 2004). However, Allison is not much concerned as I am here with accounting for Kant's apparent commitment to the knowledge of positive features of reality in itself. Instead, Allison's essay focuses more on the synthetic/analytic distinction in general and, when discussing simplicity, on Kant's specific counterarguments against Eberhard's attack. Recently, Betti (2024) has provided an elegant reconstruction of Kant's attitude towards simple beings in his mature works. While I find myself in agreement with Betti's overall stance on the topic, I nonetheless believe that his discussion overlooks some of the complications characterizing Kant's view. That is, it shall emerge from the discussion that follows that, unfortunately, Kant's attitude towards monadology is all in all less straightforward and consistent than Betti presents it — and further layers of complexity are added in the subsequent discussion of the notion of 'substantiale' (see sections 4.5 and 4.6).

can alone contain the simple, but which for us remains entirely unknowable». Something *can* contain the simple only if we take it as a thing in itself; but as such it would be unknowable for us.

The footnote attached by Kant to this passage reverts once and for all the initial impression[124]. It is worth quoting it at length, since it contains several interesting claims:

> The representation of an object as simple is a merely negative concept, which reason cannot avoid [«*der der Vernunft unvermeidlich ist*» — literally: «that is inevitable to reason»], because it alone contains the unconditioned for every composite (as a thing, not as mere form), the possibility of which is always conditioned. This concept does not, therefore, serve to extend our cognition, but merely designates a something, so far as it needs to be distinguished from objects of the senses (which all contain a composite). If I now say: that which *grounds* the possibility of the composite, and therefore alone can be conceived as not composite, is the noumenon (for it is not to be found in the sensible), I am not saying thereby: that an aggregate of *so many simple beings*, as pure objects of understanding, grounds body as appearance; but rather that nobody can have the least knowledge of whether the super-sensible which underlies that appearance as substrate is, as thing-in-itself, either composite or simple, and it is a completely erroneous view of the theory of sensible objects as mere appearances, which must be underlaid by something nonsensible, if we imagine or try to get others to imagine, that what is meant thereby is that the super-sensible substrate of matter will be divided into its monads, just as I divide matter itself; for then the *monas* (which is only the *Idea* of a not-further-conditioned condition of the composite) would be placed in space, whereupon it ceases to be a noumenon, and again becomes itself composite (8 : 209n).

The first claim one approaches is that the concept of a simple being is a merely *negative* one. What does this mean? First of all, it is clear that, in this

---

[124] In his introduction to the Italian translation of the *OaD*, La Rocca (1994) Takes note of Kant's odd claims about the simplicity of things in themselves. In order to mitigate Kant's apparent realist conclusion, he as well appeals to the footnote (8 : 209n) that I shall now discuss. He settles the issue with a condensed, but correct remark: «L'incondizionato come dimensione vuota, definita solo per negazione, si contrappone al sovrasensibile popolato di enti semplici di Eberhard» (p. 43). While I agree with such general statement, in the following I shall attempt a deeper analysis of the complex views that emerge from the text.

specific footnote, Kant is assuming that only composite beings can be given to us in experience. That said, the concept of a simple being is that of non-composite individual, which thereby cannot be given in experience.

I suggest that the most plausible reconstruction of Kant's argument in the footnote — despite some complications — is the following. We know that reason cannot avoid searching the unconditioned for every series of conditions. Every composite is always conditioned by what it is composed of. As we learn from the Second Antinomy, matter is infinitely divisible (although never given as infinitely divided), which means that no ultimate condition of composition (i.e. the unconditioned) can be found in experience. The unconditioned condition of the composite is the non-composite, i.e. the simple. This means that the concept of a simple being is therefore a product of reason in its search for the unconditioned. And it is a negative concept because it contains nothing but the abstraction from every composition. No example can be given of the result of such abstraction, and thus no objective referent can be found for <simple>. On the one hand, what is described here is pure reasons's «*unvermeidlich*» quest for the unconditioned; on the other hand, Kant reminds us that we are forced to deal with a limitation of ours, i.e. the impossibility for us to access such unconditioned condition in experience, and thus to cognize it.

Before assessing some possible difficulties for my interpretation, let us continue the analysis. Further on in the quotation, Kant connects the negative concept of simplicity to the notion of *noumenon*. A first immediate way of reading the claim that «that which *grounds* the possibility of the composite», and is therefore simple, «is the noumenon» would be the following: We are able to establish that there is one or more intelligible beings, and we know that these are the simple ground of something composite — be it some composite noumenon or phenomenon. However, Kant's words suggests an alternative reading. A simple being must be characterized as a noumenon first of all because it «is not to be found in the sensible». Simple beings cannot be given in experience, and yet transcendental idealism keeps the space open for the logical possibility of some simple *intelligible* being. The unconditioned ground of the composite remains thus a logically possible thought for us. Kant is very clear that this leaves us with no clue about the actual nature of noumena. In the rest of the passage, Kant stresses that with

the previous statement he did not mean to establish that the supersensible substrate underlying appearances is actually an aggregate of simple beings, thought by means of pure understanding alone. Rather, he maintains that «nobody can have the least knowledge of whether the super-sensible which underlies that appearance as substrate is, as thing-in-itself, either composite or simple». The passage thus confirms that Kant is *not* committed to the actuality of some intelligible being falling under the concept <simple>.

The picture is then completed by the striking specification in brackets towards the end of the footnote: the monad «is only the *Idea* of a not-further-conditioned condition of the composite». On the one hand, the claim that the monad is an *idea* seems the perfect conclusion of Kant's argument. The concept of <simple being> is produced by reason in its search for an unconditioned condition in the regressive series of composition; but no object falling under it can be found in experience, and no example of it can be taken from a purported intelligible world. This description of the monad/simple thus seems to be a typical example of Kant's notion of idea of reason, which lacks objective validity. On the other hand, at least two problems emerge from the latter claim and indeed from Kant's overall account as I reconstructed it.

First, how does the idea of a monad fit within Kant's classification of transcendental ideas? Any satisfying answer to this question would obviously require us to take issue with that classification itself, a topic that entails its own set difficulties. One of them is to establish exactly which and how many transcendental ideas of pure reason are there. To take an example from the most recent debate, Proops (2022, section 17.1) sticks to the widespread view that 'soul', 'world' and 'God' are the three ideas, but he adds that it can be said that each of them has its own '*sub-ideas*' (though he does not provide an actual table of these latter). Willaschek (2018, section 6.1), instead, prioritizes Kant's claim that the three mentioned concepts are just the three 'classes' of transcendental ideas (see A334/B391), whereas the latter ideas are actually *nine* (four under each of the first two classes and just one under the last class). None of the two authors, though, counts the monad among transcendental ideas or 'sub-ideas'. However, as Proops (2022, p. 422) notes, «absolute simplicity» is mentioned by Kant as an «idea» in a passage

from the Discipline of Pure Reason (A784/B812)[125]. It does sound legitimate then to claim that <simplicity>, or <monad> is either a transcendental or perhaps a 'non-transcendental' idea — such as for instance, it may be argued, «pure earth, pure water, pure air, etc.» (A646/B674). Considering the relevance of the notion of monad, the transcendental option appears more attractive, although things are made even more chaotic when we consider that '*simple* soul' may also be taken as a (sub-)idea.

Since it is not my aim to pursue a detailed account of Kant's classification of the ideas, I will dismiss the problem with a couple of conclusive remarks. One point worth nothing is that the difficulties just presented do not affect exclusively the notion of monad, since they concern as well, for example, the notion of 'absolute space' and, as I will later contend, that of 'substantiale'. In the *Metaphysical Foundations*, Kant explicitly writes that absolute space is «nothing more than a mere idea […] which is to serve as a rule for considering all motion therein merely as relative» (*MFNS* 4: 559-60)[126]. In this case as well, things are complicated by the fact that absolute space seems to play some role in the Antinomies, namely in the First (whereas the concept of an unconditioned ground of the composite is involved in the Second one). A major difference between the two cases, instead, is that in the *MFNS* passage above Kant does not associate the notion of absolute space with that of noumenon[127]. My last, rather trivial remark is that regardless of these systematic complications, the long footnote in *OaD* is *explicit* in calling the monad an «*Idea* of a not-further-conditioned condition of the composite»

---

[125] We find a similar claim in the Proof of the antithesis in the second antinomy (A437/B465), although that passage seems less relevant to me precisely because Kant speaks in the voice of the antithesis there.

[126] On the notion of absolute space in the *MFNS*, see for example Pollok (2001), Friedman (2013) and Massimi (2022).

[127] Massimi (2022) construes indeed the idea of absolute space in the *MFNS* against the backdrop of her interpretation of ideas of reason in general (see Massimi 2017). On her account, an idea is not to be intended as «a proxy for a noumenal entity (i.e., [in the present case,] Newton's absolute space) as some unconditioned ground required for the regulative role of reason» (Massimi 2022, p. 237). Now, I do not have specific views on the concept of absolute space at issue here, hence neither on its connection with other ideas, nor on its potential relation to some noumenal conception of space. Moreover, I have not developed a systematic account of ideas of reason in general. Nevertheless, I hope it is clear already from what I have been arguing in the present work that I am not willing to deny that (at least some) ideas involve (in complex ways) the logical possibility of noumenal entities.

which reason is bound to seek, and since it is one of the longest passages on the topic in Kant's published writings, it cannot be quickly downplayed.

A second issue, while related to the first, is worth considering on its own: How is the present argument — about the monad/simple as unavoidable idea of the unconditioned condition of the composite — compatible with Kant's position in the Second Antinomy? The first thing to note is that, whereas the dynamical (i.e. third and fourth) antinomies eventually allow for the logical possibility that the unconditioned be given outside of the series and thus in the non-sensible domain, the *mathematical* ones (first and *second*) do not contemplate that scenario (cf. A530-1/558-9). Since the mathematical antinomies concern the synthesis of the *homogenous*, if there were an unconditioned condition in these series it would need to exist within the phenomenal domain, an option that Kant eventually denies. Therefore, there would seem to be a tension between the framework of the Second Antinomy on the one hand, and the argument we are assessing here on the other. For the latter results in the claim that reason cannot avoid thinking the simple as the merely negative concept of a non-composite member that completes the series of composition[128]. Since Kant characterizes the monad here as a noumenon, it would be heterogeneous to the members of phenomenal (i.e. material) composition.

Here is a plausible solution to that tension. The Second Antinomy establishes that a *material* simple being is a *real impossibility*, and hence that the unconditioned in the regressive series of phenomenal composition ultimately cannot be given *at all*. A physical monadology (e.g. Wolff's or pre-critical Kant's one) is therefore false. Now, transcendental idealism allows at least for the logical possibility of noumena grounding phenomena. When considering the intelligible domain, we still cannot establish whether noumena are made of simple ultimate parts or not, but our reason is at least allowed to think the merely negative notion of monad/simple being as the unconditioned in a hypothetical series of composition in the purported noumenal world. On the proposed reading, phenomenal and noumenal composition would in fact give rise to two different ideas: that of a physical

[128] I will mention in the section below the further problem entitled in Kant's claim that the unconditioned may be a member of the series of conditions as well as the series itself in its totality.

monad and that of a Leibnizian monad respectively. The former is impossible, while the second alone is preserved if only as a logical possibility. Or, somehow differently, it could be the case that one and the same idea, i.e. that of 'simple being', can be specified — as it were — in two different ways, only one of which, i.e. the non-sensible Leibnizian monad, would allow us to ensure the idea of simple being the status of a merely logically possible but unavoidable thought[129].

If any of that is correct, the next question is why exactly is the idea of a simple being *unavoidable*. Here are some options. One answer could be that its *Unvermeidlichkeit* is conditional: whenever we entertain logically possible thoughts that purport to be about noumena, and more specifically about their composition, reason cannot avoid introducing the idea of monad as the not-further-conditioned condition of the composite. Alternatively, or rather in addition to that, the grounds for asserting its *Unvermeidlichkeit* may be stronger: the previous condition — namely: entertaining logically possible thoughts that purport to be about noumena — is itself in turn unavoidable (or perhaps even necessary) to reason[130]; which (on account of the previous picture) would indirectly imply that the idea of monad is equally *unvermeidlich*. A final option would be that reason needs to think an

---

[129] A problem I did not mention concerns what Kant writes in the Remark to the antithesis of the second antinomy. There Kant distinguishes the «proper signification of the word *monas*» as Leibniz intended it, from the physical monad which should more properly be called «atom» (A440-2/B 468-70). What is troublesome in that passage is Kant's claim that the Leibnizian monad is conceived as «given **immediately** as simple substance […] and not as element of the composite». Now, how can this be consistent with what we find explicitly written in the *OaD* footnote, namely that the (non-sensible) monad is the idea of the unconditioned member in a series of composition? Perhaps what Kant means to exclude in the passage at A440-2/B 468-70 is that the monad properly called be an element of *material* composition, though the text is far from clear.

[130] To make it clearer the scenario I have in mind, let me briefly illustrate a possible deflationist version of the claim that the notion of noumenon is unavoidable for reason. The (logically possible) noumenal realm *as a whole* may be characterized as a peculiar kind of unconditioned, in contrast to the phenomenal realm. For the latter — also *as a whole* — is conditioned by the cognitive subject at one end, and the non-arbitrariness of mental contents leads our reason to think that phenomena are also conditioned at the other end — precisely by things in themselves (see section 5.3). I have insisted here in taking the two realms 'as a whole', for it is not at issue the intra-domain relation among their members which may very well be all conditioned: indeed we know that this is the case for phenomena, whereas the unavailability of the PSR in the noumenal domain prevents us from answering the question of the conditionedness of any of its alleged members.

unconditioned for the regressive series of composition in phenomena no matter what, and that this leads reason to introduce the logical possibility of noumenal reality as the place where it is logically possible for the unconditioned to exist, in spite somehow of the aforementioned problem of heterogeneity[131].

Although the exact picture that emerges from Kant's treatment of the concept of <simple being> and the like is not straightforward, it is sufficient for our purposes to stress the following results. First, the notion of simple being or monad, in its content, is nothing but the *merely negative* concept of a non-composite ground of the composite. As such it is produced by reason to identify the unconditioned in the series of composition, which reason cannot avoid to seek. Second, and most importantly, there are good reasons to conclude that Kant is *not* positively committed to the existence (or even real possibility) of one or more intelligible entities falling under the concept of <simple being/monad>.

4.4 On the Analytic Truth of Monadology

I shall now turn to issue (II) presented at the outset of this chapter, namely how can we make sense of Kant's claim that monadology is intrinsically correct and what does it mean that all we know about it is derived from the definition of a world of understanding. Let us start by considering the argument advanced by Oberst (2018) in defense of the possibility of *analytical* cognition of things in themselves. He submits both a general argument and a specific one in which he considers the example of <simplicity> in order to conclude that ascription of this property to things in themselves is analytically correct. Let us first consider the latter argument and show why it does not work.

> (1) Gegenstände bestehen entweder aus unendlich vielen Teilen, aus einfachen Teilen, oder sie sind ins Unendliche teilbar.

---

[131] Another question that may be considered is whether the idea of simple being has a positive regulative role. Since it is not so important for my point here, I avoid this further complication.

(2) Aus unendlich vielen Teilen zu bestehen ist ein Widerspruch.
(3) Folglich bestehen weder Erscheinungen noch Dinge an sich aus unendlich vielen Teilen (aus (2)).
(4) Unendliche Teilbarkeit ist nur durch eine Handlung der Dekomposition möglich.
(5) Dekomposition erfordert Subjektabhängigkeit.
(6) Der Begriff eines Dings an sich schließt aus, dass Dinge an sich subjekt-abhängig sind.
(7) Folglich ist Dekomposition bei Dingen an sich nicht möglich (aus (5) und (6)).
(8) Folglich sind Dinge an sich nicht unendlich teilbar (aus (4) und (7)).
(9) Folglich bestehen Dinge an sich aus einfachen Teilen (aus (1), (3), und (8)).

(Oberst 2018, pp. 1336-7)

The main reason why Oberst's argument is not valid is because he implicitly takes for granted, without problematizing it, that *the unconditioned must actually be given*, an assumption that for Kant we are not allowed to make with regard to phenomena nor, for different reasons, for things in themselves — as I have argued in section 4.1. Oberst quotes a passage from *WRP* in support of claim (2) that being composed of infinite parts is contradictory: «For if the world, like the space and elapsed time it occupies, is given as infinite magnitude, it is then a given magnitude that can never be wholly given, which is a contradiction» (20: 289). From this passage, where Kant is discussing the mathematical antinomies, Oberst concludes that neither appearances nor things in themselves can be constituted of infinite parts[132]. However, I suggest that this conclusion is unwarranted.

First of all it must be noted that the *WRP* seems to introduce a significant change in Kant's conception of the unconditioned, which is not mentioned by Oberst. As Willaschek (2018, p. 90) notes, whereas in the *Critique* the totality of the series could count as the unconditioned, in this later text Kant takes back that possibility: «To think the concept of an absolute whole of the merely conditioned as unconditioned, involves a contradiction; the unconditioned can thus be considered only as member of the series, which

[132] «Daher können weder Erscheinungen noch Dinge an sich aus unendlich vielen Teilen bestehen» (Oberst 2018, p. 1337).

delimits the latter as ground, and is itself no consequence of another ground» (20: 287). According to the *WRP*, only a member of the series can count as the unconditioned. It is thus useful to consider the issue both from the perspective of the *Critique* and from that of the *WRP*.

In both cases, the starting point is that reason is bound to seek something unconditioned given to it. Kant then tells us that we never experience simple parts (i.e. an unconditioned member) as given in experience, and also that nothing in the world is ever given to us as already infinitely divided (and thus that we cannot access the whole series itself as unconditioned). Considering the context of the Antinomies, which Kant is summing up in the pages of the *WRP* from which Oberst takes the quotation, it seems likely to read the sentence he cites as stating that it is contradictory for an infinite whole to be «given», in the sense of 'given to a finite subject in experience'. Now, if we hold on to Kant's view in the *Critique*, there is no reason to exclude the possible hypothesis that an unconditioned itself in the noumenal domain be conceived as the totality of the series of infinite noumenal composition[133] — regardless of the fact that such totality could not be given to a finite subject in experience. The fact that this hypothesis cannot be ruled out immediately undermines premise (2) of Oberst's argument — by the lights of the *Critique*, at least. As we just saw, though, Kant's change of mind in the *WRP* about the unconditioned implies that even if the series had infinite members and, *per impossibile*, were given to us as such, it would now still not count as the unconditioned. What this suggests about things in themselves is therefore only that, *if* we assume that there actually is an unconditioned in the noumenal realm, we must conclude accordingly that noumena could not be infinitely divisible, for they would need to be composed out of simple parts, i.e. monads. Again, though, it is crucial that we cannot establish the existence of an unconditioned, not even in the noumenal domain. For all we know, noumena may thus be composed out of

[133] As I suggested earlier (see section 4.1), an unconditioned noumenon is at least a *logical* possibility, whereas its *real* possibility is theoretically undecidable, i.e. it cannot be proved nor disproved by theoretical reason. Note that, while I am discussing here the highly speculative scenario of the totality of a series located in the logically possible noumenal domain, this is not meant to contrast with the fact that Kant is mainly concerned with the question of whether a totality of a series of conditioned phenomena can be given to us in experience or not — a question that must of course be answered negatively.

infinite parts — or perhaps the picture may turn out to be entirely different in some way that we are not even able to comprehend[134].

This consideration helps us finding out which analytic truth Kant is willing to grant to monadology. Oberst concedes that analytic claims about things in themselves having certain properties, and thus falling under certain concepts, do not yet prove the real possibility of the concept at issue, e.g. <simple>. Rather, it only proves that, if things in themselves exist, they must have these properties analytically ascribed to them[135]. I will add some caveats to the latter claim, but I generally find it unproblematic that analytic propositions are only true of objects in case we are able to establish that some object actually falls under the analyzed concept. This point is relevant for our conclusion. To make it clearer, let us consider the distinction advanced by Kant in the *Only Possible Argument* between absolute and relative position (*Position* or *Setzung*)[136]. In this pre-critical essay, Kant argues against the ontological argument for the existence of God, and one of his most well-known claims to that effect is that existence is not a predicate, as many of his predecessors believed instead. In the passage that I shall discuss below,

---

[134] In his comment on Allais's (2015) book, Kreines (2016) makes a similar remark. More precisely, he correctly argues that, according to Kant, when it comes to the noumenal domain it is impossible to establish the precise nature of grounding relations — the case Kreines has in mind is that which gives rise to what he calls «regress of powers» (p. 262). Note, however, that I disagree with Kreines' view to the effect that «the regress of powers extends to properties that ground the knowable properties, more inner properties of things, unknown to us. As Kant says, appearances "must have grounds that are not appearances" (A537/B565)» (Kreines 2016, p. 262). For even if we took Kant's quoted passage at face value and conceded that appearances are grounded in things in themselves, this relation would not constitute a prosecution of the same regress of powers holding within appearances. Indeed, besides what we have established in ch. 2 with regard to the inner/outer dichotomy in the Amphiboly, it can be further observed, as Spagnesi sharply does, that the grounding relation holding among phenomena and that supposedly connecting noumena «have different subjects (phenomenal objects vs. things in themselves) and different features (spatiotemporal vs. non-spatiotemporal features)» (Spagnesi, 2024 p. 14).

[135] «Es geht nämlich nicht darum zu zeigen, dass bestimmte Eigenschaften der Dinge an sich real möglich sind, sondern nur darum, dass, wenn ein Ding an sich existiert, es bestimmte Eigenschaften haben muss. Ob Dinge an sich nun existieren oder auch nur real möglich sind, wird dadurch jedoch nicht gesagt» (Oberst 2018, p. 1336).

[136] For an interesting discussion of this distinction in relation to Kant's notion of existence, see Rosefeldt (2020). See also La Rocca (1999, pp. 74-85), who explains the notion of position as «a certain way of assuming something in the discourse [*un certo modo di assunzione di qualcosa nel discorso*]» (p. 77).

Kant's focus is quite different from our present concern about analytical propositions, for his aim here is to find an alternative notion of existence, and yet there is something interesting for us to learn from it.

When we ascribe a predicate to a subject in a sentence by means of the copula, we are 'positing' these predicates only *relatively*, that is, we establish nothing more than a logical connection between a subject and a predicate. In this way, though, we are not judging about the existence of the subject in question. By contrast, «existence is the absolute positing of a thing» (*OPA* 2: 73). If we want to posit a predicate *absolutely*, it needs first to be presupposed that the thing exists, and only then the predicate can be properly ascribed to it. Kant's example involving God makes it easier for us to understand his view:

> If I say: "God is omnipotent" all that is being thought is the logical relation between God and omnipotence, for the latter is a characteristic mark of the former. […] whether God is posited absolutely or exists, is not contained in the original assertion at all. For this reason being is correctly employed even in the case of the relations which absurdities have to each other. For example: "The God of Spinoza is subject to continuous change" (*OPA* 2: 74).

Kant's phrasing is not entirely clear on whether the predicates must be said to be absolutely posited, or rather the subject itself must be. Nevertheless, what matters to us here is that the relative positing of a predicate is a connection between subject and predicate that works independently of the actual existence of the subject of the judgement. It is thus possible to analyze a certain conception of God and assert that, given other previously established properties, God must also be thought as omnipotent. Omnipotence is here relatively posited, and our assertion could be true independently of whether God exists or not[137]. No objective referent of <God> is being thereby determined when we relatively posit God's omnipotence. It is easy to see how Kant's example of merely relative position sheds light on his later conception of analytic judgement: any analytic

[137] Kant adds that, instead of saying "God exists", it would be more proper to say that «Something existent is God», that is: «there belongs to an existing thing those predicates which, taken together, we designate by means of the expression "God"» (*OPA* 2: 74).

proposition we can obtain from a certain assumed concept of God is not by itself objectively valid unless the real possibility of such God is established[138].

That said, here is the analytical proposition I ascribe to Kant with regard to monadology:

> **Simple Being Analytic Proposition:** *If* we accept both of the following assumptions: (a) that at least one intelligible being exists and is conditioned with respect to composition, and (b) that every series of conditions in the intelligible domain is ultimately grounded in an unconditioned member of the series, then it follows *analytically* that (c) there is at least a series of composition in the intelligible domain which ends with a simple being, i.e. the «not-further-conditioned condition of the composite».

Unlike Oberst, I maintain that establishing the existence of things in themselves would not already imply the existence of monads. The notion of monad follows analytically from the concept of an intelligible domain only if the latter is such that the existence of an unconditioned member of a series is a necessary condition of that very series — which amounts to assumption (b). Kant takes this to be a common assumption among many of his predecessors, and this is why Leibniz's monadology is intrinsically correct. However, as I argued in section 4.1, the NSP (i.e. a version of Supreme Principle restricted to noumena) does not hold. Hence, even if one reads Kant as committed to

[138] I am sympathetic to Watkins and Willaschek's suggestion that analytic propositions should not by themselves already be designated as 'cognition' in the strict, proper sense. For if the latter involves the determination of an actual object, then not all analytically true propositions deserve this label (see Watkins/Willaschek 2020, p. 3209 fn. 42).

the actuality of things in themselves, this would not yet entail a commitment to the actuality of Leibnizian monads[139].

At a closer inspection, though, even if we conceded (b), it would not be sufficient that some thing in itself existed for (c) to obtain analytically. Instead, we need premise (a) formulated above, which states that 'at least one intelligible being is actual and *conditioned with respect to composition*'. Here is why. Let us suppose that (b) obtains. Even in that case, it could still be the case that an intelligible being is actual and yet there is no series of composition at all. Crucially, though, the conclusion would not obtain unless there were also a series of composition, for (b), i.e. a sort of noumenal version of the supreme principle, entails the assumption that there is a series of conditions. If (a) did not include some reference to a series of conditions, this would allow for the case that every intelligible being be such that it is not conditioned under a certain respect, e.g. composition, but only in the trivial

[139] My proposal does share some similarities with existing views on the topic such as Quarfood's and Jauernig's. Quarfood offers a brief explanation of Kant's seemingly odd claims about the simplicity of things in themselves, an explanation which, like mine, is overall deflationist. His point, though, is basically that the apparent simplicity of things in themselves is in fact nothing but a consequence of the vacuity of our discursive understanding when we abstract from sensible intuition. When we attempt to conceive of noumena by abstracting from spatiotemporal features, all we are left with is, as it were, a mirror of the structure of our discursivity. In the case of composition at issue here: «Conceptual parts can be thought of independently from the concept they compose […] so the primitive concepts can be said to be "prior to all composition", like the simples composing the substances» (Quarfood 2011, pp. 154-5). Now, the solution suggested by Quarfood is elegant, and I agree with his insistence on the shortcomings which doom the attempts of discursive understanding of establishing metaphysical principles about intelligible entities. Notwithstanding that, that does not seem to be the entirety of Kant's point. Instead, it emerged from the passages we have discussed thus far that Kant's concern here is not just with nature of those representations that are concepts. For Kant also takes the issue to involve reason and its quest for the unconditioned (and the same applies to the issue of the *substantiale* to be discussed in the next section). Jauernig (2021) too provides a valuable account, too complex to be resumed here, of Kant's apparent commitment to simplicity of noumena as well as to the Leibnizian principles criticized in Amphiboly. While I share many of her results and insights, our reconstructions differ in some crucial aspects. First of all, we diverge in that we start from different readings of Kant's distinction between negative and positive (use of the notion of) noumenon (cf. section 1.4). Also, as I have argued in section 4.1, unlike Jauernig I do not believe that we should characterize the notion of noumenon (or noumenal world) as already entailing the validity of the proposition: «Things in themselves conform to the principle that, if something conditioned exists, then its condition, and, indeed, the entire series of all of its conditions, exists as well» (Jauernig 2021, p. 350). Much less do I find Kant committed to such proposition being descriptively correct of really existing things in themselves — a conclusion that Jauernig, on the contrary, eventually suggests (see Jauernig 2021, section 6.6).

sense that it is not a part of a series of composition at all. Note that the simple is not to be defined merely as the 'non-composite', but rather more properly as the 'non-composite *ground of the composite*'[140]. Also, here is another scenario where (b), even if it were true, would lose its effect with respect to (c) in absence of (a). Imagine that the noumenal domain is populated *only* by monads, i.e. entities which, while at least potentially part of a series of composition[141], are not themselves conditioned with respect to composition. In that case, the claim that there are simple beings would not obtain in the form of an analytical consequence of (b) — although the whole reasoning would be quite pointless inasmuch as we start from the assumption that there are simple beings. In sum, we need both assumption (a) and (b) formulated above in order for the proposition (c) that 'there is at least a series of composition in the intelligible domain which ends with a simple being' to obtain analytically.

I conclude with an important remark on Oberst's more general argument. In Chapter 2 (section 2.3) I argued that Kant's strategy in the Amphiboly is to show that propositions (i1) to (i4), in spite of what may seem, are ultimately *not* analytically true precisely because they do not take into account the nature of our representational faculties. Indeed, we discover through experience that the spatiotemporal features of phenomenal reality in fact contradict these principles. Therefore they obviously do not hold for experience and neither do they hold across two worlds, so to speak, as Allais (2015) instead believes when she aims to establish that things in themselves must be the 'inner' of appearances. One may insist, however, as Oberst does,

---

[140] Although we do not find this exact expression in the long quoted footnote (*OaD* 8 : 209n), that this is how we must conceive of the simple emerges quite clearly from the passage. Now, one may nonetheless object to my point that Kant's claim that «that which *grounds* the possibility of the composite, and therefore alone can be conceived as not composite, is the noumenon» excludes from the start the hypothesis of a not-composite being which is not simultaneously a ground of the composite — and that would make my remark pointless. I believe, however, that we can easily dismiss this worry by reading Kant as concerned here with a description of simplicity, and thus of a not-composite 'properly conceived', as it were — a reading which seems to me the most charitable from a philosophical point of view.

[141] If one rejects such characterization of monads, then there would be not very much left to think about the already very abstract notion of simple being.

that those propositions are at least analytically true of things in themselves alone, and thus objectively valid under the condition that the latter exist[142].

Here as well, I contend that such interpretation is unwarranted. I already contested Wood's defense of the principle of the indiscernibles as a valid intelligible criterion (section 2.4), but let me bring here a more general point. In order to establish that the principles mentioned by Kant in the Amphiboly are at least analytically true of noumena, it is not sufficient to assume that a noumenal domain exists in general. Instead, we would need to know what it means for an *individual* intelligible being to fall under a concept whatsoever. For otherwise we are forced to work with the concept of object in general, which does not entail the specific features of intelligible objects. The acquaintance with such features is important, as we learned from the only case available to us, i.e. that of spatiotemporal objects. As the Amphiboly shows, the empty notion of object in general does not take into consideration the features of spatiality, and thus the fact that difference in space location is a decisive criterion for individuation. Since, lacking intellectual intuition, we have no clue of what makes an intelligible object an individual, it seems reasonable to conclude that we cannot know whether the principle of the identity of indiscernibles, for example, is analytically true of intelligible beings or not. For what if the criteria for the individuality of intelligible beings, which we are unaware of, ends up invalidating those principles in the way the spatiality of phenomena does? There is no way to exclude such hypothesis. Hence, although this argument cannot be found explicitly in Kant, I believe it to be a warranted outcome of what has been shown in chapters 1 and 2, and it finds special support in the structure of the Amphiboly itself.

---

[142] «Kant behauptet dort [i.e. in the Amphiboly chapter] von einigen Eigenschaften (die ich in der Einleitung genannt habe), dass Leibniz Recht gehabt hätte, wenn er sie ausschließlich von Dingen an sich behauptet hätte» (Oberst 2018, p. 1338). The properties Oberst mentions are: «dass Leibniz' Prinzip der Identität des nicht zu Unterscheidenden für sie gelte, dass bei ihnen keine Entgegensetzung von Realitäten dergestalt stattfinde, dass sie einander vermindern oder aufheben, dass Dinge an sich innere, nicht-relationale Eigenschaften haben und dass bei ihnen die Form der Materie vorhergehe» (Oberst 2018, p. 1331). A similar view is also held by Van Cleve (1999, cf. p. 145).

4.5 The Substantiale as Unconditioned

Let us now shift the attention to another concept whose mention by Kant, albeit sporadic, raises nonetheless similar worries as those connected with the just discussed notion of simplicity. Again, the ultimate goal is to achieve further insight into Kant's stance towards noumena.

The concept of <*substantiale*> (or 'substantial') is not so common in Kant's writings, and consequently largely overlooked by the scholars[143]. According to Baumgarten's definition, the substantiale is: «That in a substance in which accidents are able to inhere» (*Metaphysica* §196). In the *Metaphysik* K_3 lecture transcript we read: «Substance thought with the omission of all inhering accidents (i.e. their determination) is called the substantial» (29: 1004-5). Kant insists here as elsewhere that we cannot cognize anything of a substantiale, due to the fact that our discursive understanding can only cognize through predicates (see also, for example, *MM* 29: 904 and ML_2 28: 563). This kind of claims have led, for example, Langton (1998) and Wuerth (2014) to equate the substantiale with the notion of a noumenal substance. Indeed, Kant seems to go in that direction when he asserts, a few lines below the previous quotation, that «noumenal substances cannot be cognized» (MK_3 29: 1005); or when he claims, in a *Reflexion* quoted by Langton, that the «substantial is the thing in itself and unknown» (*Refl. 5292*). Langton and Wuerth use these results for different purposes[144], but what matters here is that both believe that Kant is thereby able to make some minimal but still positive claims concerning things in themselves. Here I shall argue against this possibility. First of all, it is worth reminding once again that in such passages it is often hard to understand whether Kant is expressing his own view or rather analyzing Baumgarten or Leibniz — the latter of which, for instance, appears

[143] Some partial exceptions can be found for example in Dyck (2014), Wuerth (2014), Messina (2021), and Proops (2021).

[144] Langton uses it in support of her reading of things in themselves as substances that have only intrinsic properties. Wuerth, instead, is concerned with establishing that the self as a noumenon is a substantiale, and thereby a noumenal substance. He attributes great importance to Kant's pre-critical view in ML_1 that «the I expresses the substantial» (28: 226), as he sees a strong continuity with Kant's critical stance. It is worth noting with Dyck that precisely the fact that the I is supposed to be given immediately in inner sense, and thus empirically, makes it less likely that in these lectures Kant viewed the I as a *noumenal* substance (cf. Dyck 2014, p. 95).

in the aforementioned page of the notes (MK_3 29: 1005). Secondly, I suggest that §§ 46-7 of the *Prolegomena*, which are devoted to the psychological idea, can shed some light on the notion of substantiale and, in connection to that, on Kant's stance towards noumena.

Before turning to those paragraphs, consider how the notion of 'substantiale' relates to the broad exegetical difficulties presented by Kant's view on the substance. The First Analogy generates serious questions for the interpreters. For example, it is not entirely clear whether any empirical object can be characterized as a substance just because it relatively persists for a finite amount of time. A further problem is to establish what exactly is necessary entailed in the category of substance. Two features in particular are often discussed, namely its *sempiternity*, and its *ultimateness* as the subject of predication[145]. The main question is thus whether or not both features are necessarily required for something to be cognized as a substance. I would like to mention Oberst (2017) here, whose recent account presents a reasonable perspective on the debate, although I suggest that it is unsatisfying on a crucial point. Oberst provides a brief but instructive historical reconstruction of Kant's view on the substance, with a special focus on the requirement of it being the ultimate substance of predication. First of all, Oberst reads the substantiale as this final subject of predication. He aims then to establish whether Kant is willing to accept that in the phenomenal world there exists something to which we can ascribe the concept of substance in what Oberst takes to be the full-blown sense, i.e. as entailing the feature of being the *substantiale*. His suggestion is then the following. Up to the first edition of the *Critique*, Kant used to tie the full-blown concept of substance to the notion of indivisibility (and thus simplicity). According to Oberst's reconstruction, then, for most of the pre-critical period Kant defended the objective validity of such notion of substance, whose instance was the physical monad[146]. From the late pre-critical period until around the *MFNS*, and thus also at the time of the first edition of the *Critique*, Kant started to deny that substance so characterized could actually be given in experience. A crucial change occurs however in the *MFNS*, for there Kant

[145] See Clusa (2024) for an overview.

[146] Oberst suggests that Kant held this view until 1769.

finally ascribes full-blown substantiality to matter, which turns out to be the ultimate subject of all that exists in space. Matter, though, is infinitely divisible as the space it fills. That testifies that Kant does not tie together anymore substantiality and indivisibility. Therefore, on Oberst's account, at the time of the *MFNS* Kant takes the substantiale to be an objectively valid concept instantiated by matter: «Das Substantiale ist dann mit der Materie identisch» (Oberst 2017, p. 17). This is the point I would like to take issue with.

It seems plausible to me that Kant takes matter to be the substance that empirically grounds all spatial bodies. Moreover, as Spagnesi (2024) insists, Kant accepts that spatial objects, i.e. parts of matter, share the substantiality of the latter as they are the ultimate subjects of certain properties. The central passage where Kant ties substance to matter is the following:

> The concept of a substance means the ultimate subject of existence, that is, that which does not itself belong in turn to the existence of another merely as a predicate. Now matter is the subject of everything that may be counted in space as belonging to the existence of things. For, aside from matter, no other subject would be thinkable except space itself, which, however, is a concept that contains nothing existent at all, but merely the necessary conditions for the external relations of possible objects of the outer senses. Thus matter, as the movable in space, is the substance therein. But all parts of matter must likewise be called substances, and thus themselves matter in turn, insofar as one can say of them that they are themselves subjects, and not merely predicates of other matters. They are themselves subjects, however, if they are movable in themselves, and thus exist in space outside their connection with other neighboring parts. Therefore, the movability belonging to matter, or any part of it, is at the same time a proof that this movable thing, and any movable part thereof, is substance (*MFNS* 4: 503).

I argue that, notwithstanding the fact that matter is in an important sense the «ultimate subject of existence», it is nonetheless *not* identical to the substantiale, since the latter concept is defined by Kant in such a way that it cannot be applied to any empirical being, including matter.

However one may read Kant's concept of matter in the *MFNS*, it is relatively uncontroversial that it has some fundamental determinations

which characterizes it as such, be it impenetrability and/or movability[147]. Crucially, such predicates are essential to matter's very role as the ground of material objects, for the latter objects *inherit* to some extent the attributes of matter[148]. Hence only if we take matter *together with its fundamental predicates* can we identify it as the ultimate subject of all that exists in space and is not in turn a predicate of another being. More than that, in abstraction from these predicates, *we could not even conceive matter as such in the first place*. This is the reason why neither matter, nor any determinate individual we can conceive of, can be said to be identical with the substantiale. As the passages quoted above from the *Lectures on Metaphysic*s show, the concept of <substantiale>, since it aims to designate that in a substance which grounds its accidents, is thus an abstraction from any predicate whatsoever. One may even say, paralleling what Kant claims about <simplicity>, that <substantiale> is nothing but a merely negative concept. Lacking any predicate, an instance of the substantiale would be for us *indistinguishable* from anything else. This is consistent with the first category of relation being <substance and accident>, for no substance is given to us in experience if not together with at least some accidents inhering in it. Therefore, matter is *substance* but not the *substantiale*, a concept which indeed is not mentioned at all in the *MFNS*.

To be sure, this is an epistemological point concerning our discursive faculties. Nevertheless, its related metaphysical counterpart comes in a regulative fashion: *Reason is bound to seek the substantiale*, i.e. the idea of the unconditioned member in the series of subjects and predicates. Yet, as usual, no object of experience, and consequently no object we can cognize at all can properly instantiate the substantiale. Assuming that the substantiale exists equates assuming the unconditioned, and this assumption must be regulative and not constitutive.

In defense of the reading just outlined, I shall now consider §§ 46-7 of the *Prolegomena*. Oberst understandably dismisses the passage therein where Kant mentions the substantiale, for he is not able to integrate them in his

---

[147] For different perspectives on what should count as a fundamental determination of matter, compare for example Watkins (1998), Friedman (2001) and McLear (2018).

[148] For instance, since spatial objects are parts of matter, «they are movable in themselves» — as Kant writes in the passage just quoted (*MFNS* 4: 503).

account[149]. In contrast, I take the passage to be especially indicative of Kant's stance towards the substantiale. Here are the main points of Kant's argument in those paragraphs. He starts §46 by mentioning the 'substantiale' as a well-known concept which signifies «the true subject», i.e. «that which remains after all accidents (as predicates) have been removed» (4: 333). As in the lecture transcripts, this absolute subject is said to be unknown to us. Nevertheless, he insists that this impossibility is a mere consequence of the fact that our understanding always cognizes by means of concepts, and that all concepts, as possible predicates of a subject, are accidents of a substance. Since there is no cognition without concepts, or accidents, there can be no cognition of a substantiale. What is most striking in the passage, though, is that Kant seems nonetheless to grant the substantiale an heuristic role. Here is the full passage:

> But it needs to be said that human understanding is not to be blamed because it does not know the substantial in things, i.e., cannot determine it by itself, but rather because it wants to cognize determinately, like an object that is given, what is only an idea. 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). But from this it follows that we should take nothing that we can attain for a final subject, and that the substantial itself could never be thought by our ever-so-deeply penetrating understanding, even if the whole of nature were laid bare before it; for the specific nature of our understanding consists in thinking everything discursively, i.e., through concepts, hence through mere predicates, among which the absolute subject must therefore always be absent (*Prol.* 4: 333).

The substantiale is nothing more than the *idea* of an absolute, final subject. Kant's maintains that human understanding «is not to be blamed» simply

---

[149] «Über die Gründe dafür kann nur spekuliert werden, möglicherweise handelt es sich um ein Versehen» (Oberst 2017, p. 7 fn. 17). Yet the unknowability of the substantiale, conceived as an abstraction from all predicates, can also be found in lecture transcripts (quoted above), which are posterior to the *MFNS*, as well as in the *WRP*. In the latter we read that «the substantial […] remains behind after I have taken away all the accidents that inhere in it, but absolutely cannot be known any further at all, since the accidents were precisely that whereby I was able to know its nature» (*WRP* 20: 270).

because of its discursivity, «but rather because it wants to cognize determinately, like an object that is given, what is only an idea». The substantiale exceeds by definition the limits that, according to Kant, characterize our discursive understanding (and which have been discussed in Chapter 1). Nevertheless, it is reason itself that generates this idea in its quest for the unconditioned and prescribes us to search for a further subject ever again until we reach the substantiale, which however is not available to us. Because of its independence from the accidents that it grounds, it is reasonable to maintain that the substantiale would be the unconditioned member in a series of inherence. It can be said that the notion of substantiale expresses at the same time, on the one hand the limits of our discursivity, and on the other the need for reason to seek ultimate grounds.

The example of the self introduced by Kant here fits quite well the results of our previous discussion of simplicity. As he reminds us here, according to his diagnosis, it might seem to us that we have an intuition of a given substantiale through apperceptive self-consciousness, namely the I (or thinking subject) as the ultimate subject of all my inner states. If that were the case, suggests Kant, it would appear that the substantiale is not a mere idea, but something given to us after all. As we know, and as Kant makes immediately clear, this however is *not* the case. The «thinking self (the soul), as the ultimate subject of thinking» (4: 334), while not given to experience, is for us an idea (which serves thereby as a regulative principle). Kant means of course here the idea of the soul. The ultimate subject, concedes Kant, «may now indeed be called substance»: but if persistence «cannot be proven of it», then it cannot be the schematized concept of substance that has a constitutive role in our cognition. And indeed, as expected, he firmly maintains that we are not able to prove persistence from this concept of thinking being depicted as an ultimate subject — as we also learn from the Paralogisms. This impossibility is summed up by Kant with the following important remark: «Persistence, however, can never be proven from the concept of a substance as a thing in itself». Now, it seems clear from the context of the passage that the expression «substance as a thing in itself» refers to the mere idea of the thinking being as absolute subject, which is in turn an example of the negative concept of substantiale.

Prolegomena §§ 46-7 thus suggest that the substantiale may be called a «substance as a thing in itself», but at the same time warns us against taking its existence as a matter of fact. It is the concept of a thing in itself because it is not an empirical concept, nor a schematized category (and thus differs from the schematized category of <substance and accident>). It designates something that we cannot experience nor even grasp through our discursive understanding. Perhaps an intuitive intellect would grasp it immediately, because it needs not cognize by means of predicates in judgment. Such intellect might thus be able to distinguish among different individual *substantialia*: but this is nothing more than a hypothesis, involving empty concepts whose exact meaning we cannot even grasp. Here we get a first response to the puzzling claims of the *Lectures on Metaphysics*: Kant is not committed to things in themselves falling under the concept of <substantiale>. Another result is that we can put forward an analytical proposition about noumena the way we did with the concept of <simple being>. In the present case:

> **Substantiale Analytic Proposition:** *If* we accept both of the following assumptions: (a) that at least one intelligible being exists and is conditioned with respect to inherence, and (b) that every series of conditions in the intelligible domain is ultimately grounded in an unconditioned member of the series, then it follows *analytically* that (c) there is at least a series of inherence in the intelligible domain which ends with a substantiale, i.e. the unconditioned condition of its attributes[150].

I would like now to discuss some difficulties in my account, and I shall start to do so by engaging with a few considerations advanced by Proops which fit the purpose. Proops provides one of the few accounts of the Transcendental Dialectic which brings some attention, if only briefly, to the notion of substantiale. Now, consider the following remarks:

---

[150] Claim (b) is the same that has been already discussed in occasion of the 'Simple Being Analytic Proposition'. As for (a), it is obviously different in the two cases, since inherence is involved in the present case instead of composition, but I take similar remarks to apply here as those advanced in the previous section.

> although Kant does allow that we are conscious in pure apperception of the *substantiale*, this is not something that we are held to be conscious of as exemplifying the unschematized category of substance. It is rather the referent of the I thought of merely as a something that bears accidents; it is the self considered in abstraction from its accidents (29: 904). To think of the I in this way is emphatically *not* to think of it as exemplifying the unschematized concept of substance, for the latter concept includes the modal notion of being *incapable* of inhering in another thing. (Proops 2022, p. 135).

According to Proops, in pure apperception I am aware of myself as an instance of <substantiale>, a concept which on his reading, however, does not bring with it any metaphysical consequence about my self as an object. Using the terminology I employed in Chapter 3, the substantiale would seem for Proops to express a certain condition of self-ascription of my thoughts. I have instead insisted on another reading of the substantiale in relation to the self: Since in pure apperception I think myself as instantiating the identical correlate to which my thoughts are to be ascribed, it merely *seems* to me that I am an instance of <substantiale>. For, differently from Proops, I take the latter to express the unconditioned subject of inherence. On my account, then, the substantiale characterizes that which the 'rationalist' — who assumes the unconditioned to be actually given and ignores the unique features of pure apperception — believes to be an object given in inner sense. Substantiale is then the object of reasons's illusion, and not the correct-but-metaphysically-sterile way we are aware of ourselves in pure apperception.

To be sure, some texts seem to speak in favor of Proops' proposal. For instance, this appears to be Kant's view in the previously quoted passage of *WRP* where he clarifies the meaning of 'substantiale':

> Of the self in the first sense (the subject of apperception), the logical self as a priori representation, it is absolutely impossible to know anything further as to what sort of being it is, or what its natural constitution may be; it is like the substantial [*es ist gleichsam, wie das Substanziale*], which remains behind after I have taken away all the accidents that inhere in it, but absolutely cannot be known any further at all, since the accidents

were precisely that whereby I was able to know its nature» (*WRP* 20: 270).

The whole passage is puzzling on various levels, because it appears to identify the «logical self» with the existing being that apperceives. As I mentioned at the end of Chapter 3, I take his claim below in the text that «the logical self does indeed point to the subject as it is in itself […] as pure spontaneity» (*WRP* 20: 270-1) to suggest a connection and yet at the same time a distinction between the 'I' of the 'I think' and whatever existing being that apperceives. In light of this remark, one possible reading of the long quoted passage is that the substantiale may indeed purport to refer to the subject to which the logical self 'points'. Note, however, that it is also possible to take the passage to establish a link between the representation 'I' and that of substantiale without thereby conceding that the 'I' of the 'I think' is an instance of <substantiale>. Indeed, what the expression «it is like [*es ist gleichsam, wie*]» seems to suggest is that Kant is drawing here only an *analogy* between the representation 'I' and that of substantiale, to the extent that both represent a content to which no predicates can be objectively ascribed.

Proops himself cites a page from Kant's *Lectures on Metaphysics* (i.e. *MM* 29: 904) in support of his claim. It suffices to note that the passage is confusing on its own in that it is particularly hard to establish therein when Kant is expressing his own views, and when he is speaking on behalf of Baumgarten instead. For example, the claim that «I am the ultimate subject and I *cognize* myself without accidents [italics mine]» can hardly represent Kant's own position. Indeed, the impossibility to be aware of ourselves as a substantiale is confirmed by a passage of the B-Paralogisms:

> I *confuse* the possible abstraction from my empirically determined existence with the supposed consciousness of a separate possible existence of my thinking Self, and *believe that I cognize* what is *substantial* in me as a transcendental subject, since I have in thought merely the unity of consciousness that grounds everything determinate as the mere form of cognition (B427, italics mine).

What we read here is perfectly in line with Kant's much discussed point in the Paralogisms that pure apperception is not to be equated with the

awareness of an actual substance, only that here he speaks of the «*Substantiale in mir*».

Central in my account is that, against Proops' *denial* that the <substantiale> «includes the modal notion of being *incapable* of inhering in another thing», I maintain that the substantiale is an unconditioned final subject, as it results from the equivalence explicitly drawn by Kant between a «*letztes Subjekt*» and the substantiale (cf. *Prol.* 4: 333). Now, in the previously quoted passage, Proops takes trouble to stress that, «although Kant does allow that we are conscious in pure apperception of the *substantiale*, this is not something that we are held to be conscious of as exemplifying the unschematized category of substance» (Proops 2022, p. 135). It must be clear, though, that my account of the substantiale does not need to concede either that the latter 'exemplifies' the unschematized category of substance — if «exemplifying» here means for something to be an actual individual falling under the unschematized category of substance. The reason for this is that we can never provide an objective referent for the substantiale in the first place. In fact, an instance of <substantiale> could never be a possible object of experience because we can never cognize a substance in abstraction from its predicates. We thus cannot cognize and point at any concrete instance of substantiale that would in turn serve as an 'example' of the category of substance, be it the unschematized or the schematized one.

A further noteworthy point is brought up again by Proops. In discussing Kant's account of the transcendental illusion involved in the idea of the soul, Proops compares Kant's stances in the *Prolegomena* and in the *Critique*. The issue is whether or not the idea of the soul involves a series of conditions. According to the *Prolegomena* the answer would seem to be yes; according to the *Critique*, no. Having noted this inconsistency, Proops proposes the following compromise:

> what goes (or, at any rate, should go) is the idea of a series of conditions, while what stays (or should stay) is the idea of a subject that is not (and cannot be) a predicate of something else, and which is, in that sense, absolutely unconditioned. Kant is not entirely consistent on this point, but this seems like the best way to try to make him consistent (Proops 2022, p. 132).

In a footnote he stresses some further inconsistency in Kant's position within the *Critique* as well: «The account of transcendental illusion in the introduction to the Dialectic and in the chapter on Ideas is not entirely consistent with Kant's remarks in the Antinomies. For in the two former places he maintains that all ideas give rise to regressive series of conditions» (Proops 2022, p. 132 fn. 14). All this is relevant for my formulation of the analytic proposition concerning the substantiale in general, for there I spoke of a *series* of inherence. Given the internal inconsistencies in Kant's text, though, I leave it open how it may be called what I have designated with the term 'series', namely the process whereby any relative substance can in turn be taken as the accident of (and thus as conditioned by) a more fundamental substance.

Let us consider a final complication. When introducing the System of Cosmological Ideas at the beginning of the Antinomy chapter, Kant explains why the Third Antinomy involves precisely the category of causality, more specifically, a regressive series of causes and effects. After excluding that a series could concern the category of <substance and accident>, he briefly considers the thinner notion of *substantiale*:

> What might still seem to be an idea of transcendental reason here would be the concept of the **substantial**. Only since this signifies nothing other than the concept of a subsisting object in general, insofar as one thinks in it merely the transcendental subject without any predicates, but here only the unconditioned in a series of appearances is under discussion, it is clear that the substantial cannot constitute a member of that (A414/B441).

Kant's claim that the substantiale in general is not an idea contrasts with the view I proposed. In fact, though, it contrasts with Kant's own words in the *Prolegomena*, since there, as we have already seen, he explicitly characterizes it in precisely these terms. Here is the passage again: «human understanding is not to be blamed because it does not know the *substantial* in things, i.e., cannot determine it by itself, but rather because it wants to cognize determinately, like an object that is given, what is only an *idea*» (*Prol.* 4: 333, italics mine). This inconsistency cannot be dismissed by simply supposing that in the latter passage Kant is concerned with the soul alone, for he

mentions instead the «*substantial* in things», and thus in general. Now, the quotation from the *Critique* can be read as to mean the following: The reason why the substantiale cannot be «the unconditioned in a series of appearances» is because it is a mere abstraction and/or because, if there existed an instance of such abstraction, this could not be an appearance, but rather only an intelligible object of intellectual intuition. This answer, though, is not fully satisfying, for one may further raise the following concern. The idea of an uncaused cause, like the other transcendental ideas, is also eventually not given as an object of possible experience, and yet this is an *outcome* of Kant's argument in the Antinomies, not his starting assumption. So why does Kant instead exclude already *from the beginning* that the substantiale be an idea of reason? Unfortunately, I have to admit that Kant's systematic remark is not entirely clear to me.

4.6 Monadology, Substantiality and the Soul

The analysis of the notion of substantiale has provided interesting results. On the negative side, we learned that the substantiale cannot be given as such in experience, and thus that not even matter is identical to it. Neither does Kant take to have established its objective validity in the noumenal domain. The substantiale is rather a negative concept, an abstraction which we are not fully capable to grasp because it exceeds by definition the limits of our discursive understanding. There is also a positive side of the story, even though it brings with it the just discussed difficulties in making Kant's texts entirely consistent. We found textual evidence of Kant entertaining the view that the substantiale is another case of unconditioned condition that reason is bound to seek, namely the unconditioned subject of inherence.

Things get especially chaotic when we attempt to figure out the connection between the idea of substantiale and that of the soul. The latter seems to be a special case of the former; or, put differently, that of being a substantiale is an alleged feature of the soul. The picture becomes even more confusing if we consider the relation between these two notions and that of monad, or simple being. Connected to them is finally also the concept of 'inner' (or intrinsic properties). Something all these metaphysical concepts have in common is

that they play a central role in Leibniz's monadology, as they all converge in the notion of the monadic soul. Indeed, according to Kant's interpretation of Leibniz, inner sense is taken by the latter to be a privileged access to a monad, namely to that which we ourselves are:

> what can I think of as inner accidents except for those which my inner sense offers me? — namely that which is either itself **thinking** or which is analogous to one. […] taking away in thought everything that might signify outer relation, thus even **composition**, Leibniz made out of all substances, even the constituents of matter, simple subjects gifted with powers of representation, in a word, monads (A266/B322).

In the *Metaphysical Foundations*, Kant provides a clarification of the notion of 'inertia' which makes it in turn clearer why Leibniz conceived the monad the way he did:

> The inertia of matter is, and means, nothing else than its **lifelessness**, as matter in itself. **Life** is the faculty of a **substance** to determine itself to act from an **internal principle**, of a finite substance to change, and of a material substance [to determine itself] to motion or rest, as change of its state. Now we know no other internal principle in a substance for changing its state except **desiring**, and no other internal activity at all except **thinking**, together with that which depends on it, the feeling of pleasure or displeasure, and **desire** or willing. But these actions and grounds of determination in no way belong to representations of the outer senses, and so neither [do they belong] to the determinations of matter as matter. Hence all matter, as such, is **lifeless** (*MFNS* 4: 544).

Given Kant's denial of the possibility to cognize ourselves as a simple substantial soul through inner sense, this *MFNS* passage is best read as a sketch of empirical psychology, where the idea of soul may indeed play a central heuristic role. Yet what is relevant for us here is that the two previous quotations combined show that, on Kant's reconstruction, Leibniz took faculties such as desire and thinking to prove our nature of simple beings. In fact, on Kant's reading, Leibniz could find no other way to directly prove the objective reality of the monad.

Now, none of the concepts mentioned above (i.e. <soul>, <inner>, <monad/simple>, <substantiale>), which are central to rationalist metaphysics, can however be encountered in experience[151]. Furthermore, our investigation has shown that we can neither establish their objective validity by providing an actual example in the noumenal domain. That is: *Kant is not committed to a noumenal monadology.* Nonetheless, Kant's account of reason's intrinsic propensity towards metaphysics allows him to reconsider these intertwined concepts in a new light by ascribing them a regulative status. He refrains from dismissing such concepts entirely, even when this threatens to undermine the overall consistency of the system, as in the case of the <simple> and the <substantiale>.

The considerations advanced in the present chapter shed some further light on the previously discussed issue of the ontological status of the self. First of all, as stressed by some scholars, we face a major disparity in the possibility to apply the category of substance and accident to spatial things versus our inner states. Outer objects exist in both space and time, and that which exists in space can be said to persist through time. We have seen that in the *MFNS* Kant characterizes matter as the substance which, together with its accidents, grounds all that exists in space. The structure of matter ensures the validity of the category of <substance and accident> as *constitutive* for spatiotemporal objects. The only form of inner sense, instead, is time[152]. The manifold of inner sense thus lacks the structure of matter and cannot provide us with something that persists throughout time. Hence the category of <substance and accident> has no constitutive role there, which entails that it cannot be legitimately applied to the soul as the purported object of inner sense. The situation is different with respect to the notion of <substantiale>: here there is no proper disparity between the two types of phenomena. For, as we have seen, no instance of <substantiale> can be presented in experience *at all*, be it inner or outer — nor anywhere else. In this case, then, it is not because of some structural defectiveness of the manifold of inner

---

[151] As already noted (section 2.2.), <inner> can nonetheless be employed in experience as a *relative* concept, that meaning that some properties can be comparatively 'inner'.

[152] On Kant's diagnosis, the non-spatiality of inner states is a «natural and popular» argument which *mistakenly* leads «even the commonest understanding» (A357-8) to «except [the soul] from the perishability to which matter is always subjected» (A357).

sense if the substantiale cannot be given therein, but rather as consequence of understanding's discursivity[153].

Since matter is a substance properly intended, and given that the actuality (or even real possibility) of an unconditioned cannot be proved, an instance of <substantiale> seems to match the definition of unconditioned better than an instance of <substance>. It is therefore tempting to read the concept of <substantiale> — instead of <substance> — as the main apparent feature of the soul, i.e. of the unconditioned ground of my inner states. To be sure, this proposal finds almost no terminological validation in the *Critique*, where the substantive '*Substantiale*' is almost entirely absent. Furthermore, it would need to be investigated more in detail how to make Kant's characterization of the soul as a regulative idea consistent with his views on the substantiale, something I will not pursue here. On the other hand, though, the concept of <substantiale> nicely squares with the hypothetical attributes of the soul discussed in the Paralogisms.

I conclude by showing how these considerations allow us to rethink the supposed *noumenality* of the soul, i.e. its existence as an intelligible being. In the previous chapter, I have stressed that Kant denies in the Paralogisms that we can achieve any sort of metaphysical determination of the soul. Accordingly, I have argued, among other things, for the negative but crucial point that the kind of self-awareness at work in Kant's pure apperception is *not* an awareness of myself as noumenon. Nevertheless, I would like to note here that the possibility of a noumenal soul is entertained by Kant for several reasons, among which at least the following could be listed. First, it is introduced to refute materialism[154], whose consequences would otherwise be

---

[153] With these limited remarks I do not mean to engage with the dispute on the parity/disparity of inner and outer sense in its entirety. For an overview on that, see Kraus (2019).

[154] On Kant's refutation of materialism about the soul, see Kraus (2021). Consider that Kant nevertheless admits (e.g. at A359-60) that we *cannot disprove* the possibility according to which the noumenal ground of matter and that of our inner states are the *same*. Ameriks discusses at length the complex issue of noumenal immateriality (2000, ch. 2) and suggests that Kant's openness to such possibility should not be taken at face value. It seems to me that, given the restrictions on our access to noumena, Kant would be correct to maintain the possibility that inner and outer appearances have the same non-sensible ground. At the same time, though, such hypothetical identity would eventually turn out to be hard to square with Kant's original presuppositions about noumenal affection — which I shall discuss in the next chapter.

devastating from the moral standpoint. Secondly, and in close connection to that, the idea of an intelligible, non-phenomenal soul plays a notoriously central role in Kant's practical philosophy. A third reason, whose overlooked importance I would like to stress here, is that it provides a metaphysical framework that warrants our investigation of inner states and faculties. Indeed, the logical possibility of the noumenal soul as an instance of <substantiale> opens in turn the possibility that faculties have an unconditioned ground. Since a power is described by Kant as the relation of a subject to its accidents, more precisely the capacity of a subject to bring about accidents and ground them[155] (in this context: soul's inner states), I find it plausible to claim that, for Kant, the absolute impossibility of a substantiale would have compromised the very legitimacy of employing properly the notion of 'power' or 'faculty' to investigate our mind.

Now, crucially, it must be noted that there is in principle no theoretical reason to exclude that the noumenal ground of our faculties and inner states be *multiple* intelligible entities acting jointly, and thus perhaps also a *set of substantialia*[156]. It might as well be that the soul is actually nothing but an accident of some other intelligible being[157]; or perhaps even the byproduct of some inscrutable noumenal interaction among elements of noumenal reality — whatever this sentence might mean. Note as well that the necessity of the relation between a substance, its accidents, and the powers through which the former brings about the latter are established in the First Analogy and

---

[155] Kant attacks Wolff's view according to which, in Kant's reported words, «the soul itself is a basic power which represents the universe». In order to support this criticism, then, Kant makes the following ontological claim: «Power [Kraft] is not what contains in itself the ground of the actual representation, but rather the relation <*respectus*> of the substance to the accident, insofar as the ground of the actual representation is contained in it» (*ML*$_1$ 28: 261). In the *OaD*, he writes: «Force [*Die Kraft*] is not that which contains the ground of the existence of accidents (for substance contains that); it is rather the concept of the mere relation of substance to the latter, *insofar* as it contains their ground» (8: 224n).

[156] *Contra* Wuerth (2014), who, as already mentioned, is one of the most strenuous defenders of the opposite view.

[157] Similar options are contemplated e.g. by Ameriks (2000, p. 69) and Marshall (2010). The latter attributes to Kant a «effect-relative compositional approach to the self». While not excluding that «we might in fact be one or more substances», he rightfully maintains that «we have no a priori warrant for regarding ourselves as such» (Marshall 2010, p. 10). That said, both he and Ameriks find Kant more deeply committed to the existence of the soul so characterized than I am willing to concede here.

thus holds only within the domain of possible experience. The point is therefore that we humans have no tools to resolve the issue. All this, however, does not count against the proposed view, but rather supports it. For I maintain that we are only able to establish the *logical* possibility of the soul as an unconditioned substantiale, and that we need no more than that to pursue our investigations of the faculties of the mind[158]. In conclusion, theoretical reason cannot and does not need to establish the existence of a substantial soul.

***

In the first section of this chapter I argued that Kant is *not* positively committed to the existence of the unconditioned on theoretical grounds, as he does not hold a noumenal version of the supreme principle (NSP) to be true. In particular, we have seen that Kant's diagnoses of transcendental realism's mistake does not prove a commitment to the NSP. The remainder of the chapter was instead devolved to the suspect, arising from a series of ambiguous passages in Kant's works, that Kant's noumenal reality may have retained the core features of Leibniz's (and others predecessors') monadology. This investigation has brought a negative result to the effect that Kant is *not* committed to the existence of an intelligible world populated by monads. The analytical propositions about noumena involving the concepts of <simple being> and <substantiale> rely on possible, but ultimately unwarranted assumptions. Along the way, however, we have also witnessed Kant's strenuous efforts to come to terms with the metaphysical doctrines of his predecessors. The notions, or perhaps better 'ideas' of <simple> and <substantiale> play a complex role in Kant's system as negative concepts of an unconditioned that reason is inevitably bound to seek. In particular, we have seen that the elusive notion of <substantiale>, an unconditioned subject of inherence, cannot be properly grasped by our discursive understanding, and cannot be instantiated by matter, nor by any object given to us in experience. At the same time, though, we cannot establish whether an intelligible being exists that falls under such concept.

---

[158] One may add that this logical possibility subsequently becomes an object of practical cognition on moral grounds, in case the substantiality of our soul be required for our moral responsibility as individual persons before God's judgement.

We have finally seen that the notions of <substantiale>, <simple being>, along with that of <inner> merge together in the Leibnizian concept of a monadic soul. The idea of a non-empirical <substantiale> underlying our inner states is no more than a logical possibility, which nonetheless plays several crucial functions in Kant's system. Overall, I hope this chapter has proved that it is possible to take seriously Kant's positive reappraisal of 'Leibnizian' metaphysics while at the same time defending a deflationist reading of transcendental idealism.

5
Transcendental Idealism Pushed to its Limit: The Existence of Noumenal Reality

5.1 The Import of my Deflationist Account

In the course of this work I have explicitly pursued an overall deflationist interpretation of Kant's stance towards noumena. Special attention was brought to the impossibility for us discursive beings to single out individual non-sensible entities. We cannot be aware of what are the criteria for individuating concrete things in abstraction from sensible conditions — or rather, more radically, *whether* there are any such criteria. Hence the hypothesis of an individual non-sensible thing falling under a concept, while logically possible, is barely graspable for us. Accordingly, claiming to have established the actuality or real possibility of noumena amounts to a positive and unwarranted use of the notion of noumenon. Furthermore, we have seen time and again that no metaphysical principle is available for us to establish on theoretical, objective grounds necessary connections between empirical objects and noumena, nor among noumena alone. Even the idea that we can make truthful analytical claims about noumena has been challenged and qualified. Analytical truths of this kind should often more properly be said to hold not of purported non-sensible *things*, but rather of the sort of hypostatized concepts which we are left with — due to the shortcomings of our discursive understanding — when we abstract from space and time. Also, when discussing the notions of <simple being> and <substantiale> in the previous chapter, I have put forward analytical propositions involving the non-sensible domain which however hold only if we take them to entail specific, indemonstrable assumptions (e.g. the validity of the NSP).

Importantly, though, this does not disprove the existence of non-sensible beings. Not only the notion of individual noumena, and more generally of a noumenal domain or layer of reality are logically possible; Kant also believes that his transcendental idealism secures that the *real* possibility of noumena is theoretically undecidable, meaning as well that their real *im*possibility is

indemonstrable[159]. Let us refer to the conjunction of these two claims concerning the possibility of noumena as the Kantian thesis of the 'logical possibility+', or 'LP+' of noumena. This is of utmost importance for Kant's moral philosophy, which makes objective, though only practically grounded claims about noumena. Moreover, as we have seen in the previous chapter, the LP+ of noumena legitimizes pure reason's never-ending quest for an unconditioned condition for certain series of conditions.

A brief clarification on the LP+ of noumena is in need. One may object that the assessment of the theoretical indemonstrability of noumena's real *im*possibility included in what I called 'LP+' amounts already to admitting the *real* possibility of noumena. In fact, the objection would go, such indemonstrability implies that it is in some sense *possible* for noumena to be really possible. What kind of possibility is then at issue here? And most importantly, is it not such possibility the possibility *of noumena*, after all? While I do not have a straightforward Kantian answer to the first issue, the second question must be answered with a simple *no*. When Kant assesses the theoretical indemonstrability of noumena's real impossibility, he is not establishing the real possibility of noumena, but rather the mere possibility of a *modal fact* concerning noumena. More properly, what is possible are two contradictory modal facts about them, namely that they are really possible and that they are really impossible. The point is thus that, due to the limitations of our faculties, we are not in the position to establish on theoretical grounds which of these two contradictories is true. Had we established the real possibility of noumena, we would have thereby excluded their real impossibility. Crucially, however, this is not the case: both alternatives, i.e. both modal facts are (in some sense) possible. All we know is that we have no cognitive access to either the real possibility or impossibility of noumena. And yet this is not a minor achievement for Kant, since it is enough to silence materialism and open the space for the demands of pure practical reason (cf. BXXIX-XXX).

Now, the deflationist core of my interpretation of Kant's transcendental idealism can be summed up in the claim that the LP+ of noumena is all that

---

[159] As I mentioned in note 112 in section 4.1, Kant takes this to be a major achievement of the Transcendental Dialectic.

we have and all that Kant needs, which means that we have no clue about their actuality or even real possibility. However, I want to reconsider my position here by distinguishing two claims which are not equally plausible as an interpretation of Kant's view. (a) The first claim is that we cannot establish (nor deny) on objective theoretical grounds the existence or the real possibility of any concrete *individual* non-sensible being: I take this, and related claims, to represent Kant's own view. Properly understood, objectivity and individuation are essentially tied to the spatiotemporal domain. Talking about individual noumena or discussing whether things have or not in themselves certain properties is a mere manner of speaking, since Kant remains neutral on whether the very notion of individuality ultimately makes sense outside the field of possible experience. Individual non-sensible beings are a mere LP+.

It is important to see, though, that claim (a) does not yet immediately exclude that we might be somehow able to demonstrate the existence of *non-sensible reality in general*, even if this requires us to conceive it as a sort of indistinguishable magma, as it were, which somehow partially provides, by grounding it, the matter of experience. A separate issue is therefore whether it is plausible to ascribe to Kant the following claim (b): We cannot establish (nor deny) on objective theoretical grounds the existence nor the real possibility of *non-sensible reality in general*. I believe that, while (b) *should* be Kant's view, eventually it *is not*.

Admittedly, the latter statement is exegetically unappealing, since an interpretation of Kant that is able to conciliate his theoretical commitment to non-sensible reality with the limitations of discursive cognition appears to be preferable. Yet I believe that if we loosen Kant's limitations on discursive knowledge to accommodate his commitment to non-sensible reality, his own account of such limitations becomes less consistent and philosophically less plausible. I will not elaborate further on this debatable point. Instead, in the remainder of this work I shall first of all consider Kant's most explicit argument (or so it seems) for the existence of non-sensible reality to the effect that 'appearance' implies 'something that appears', in order to focus on a second worry: I suspect that Kant's commitment to the existence of non-

sensible reality is in fact not argued for, but rather simply presupposed[160] (and it is the core of a set of unwarranted initial assumptions concerning affection and human faculties). While the LP+ of non-sensible reality appears sufficient for most of Kant's purposes, it does not seem to be enough for Kant's account of the origin of experience.

5.2 The 'Absurdity Claim'

In several passages Kant proposes some version of what may at first glance appear to be a sound argument for the existence of things in themselves. In the following, I will show that it is not, for noumenal reality is therein *presupposed* rather than argued for. Let us start by considering what is perhaps the most famous of those passages. In the B-Preface Kant writes:

> we can have cognition of no object as a thing in itself [*Gegenstande als Dinge an sich selbst*], but only insofar as it is an object of sensible intuition, i.e. as an appearance; from which follows the limitation of all even possible speculative cognition of reason to mere objects of **experience**. Yet the reservation must also be well noted, that even if we cannot **cognize** these same objects as things in themselves, we at least must be able to **think** them as things in themselves [*daß wir eben dieselben Gegenstände auch als Dinge an sich selbst, wenn gleich nicht **erkennen**, doch wenigstens müssen **denken** können*]. For otherwise there would follow the

[160] This conclusion is far from new. Not only post-Kantian German philosophers, such as Hegel, have criticized Kant on this point. Some contemporary Kantian scholars as well have noted that Kant's commitment to things in themselves is a core presupposition of his system — though some are more critical (e.g. Strawson 1966) than others (e.g Gardner 1999, ch. 8, and Jauernig 2021, section 5.10.2) towards this and other assumptions made by Kant at the outset of the critical enterprise.

> absurd proposition that there is an appearance without anything that appears [*daß Erscheinung ohne etwas wäre, was da erscheint*] (BXXVI-VII)[161].

We find here what I shall call the 'Absurdity Claim', that is, the claim that it would be absurd for there to be an *appearance* without *something that appears* in it. Kant maintains that if this «absurd proposition» is to be avoided, then we must «**think**» (though we cannot cognize) things in themselves. Or rather, if we are to remain more faithful to the text: we must think of objects of experience in general as things whose constitution is independent from sensible intuition. Given the reservations about strict identification of appearances and things in themselves (see note 161), and given the impossibility to count individual noumena, we can further rephrase the latter point as follows: we must think a level of reality which in itself is not sensible.

A first question is what the expression 'something that appears' supposedly refers to. There is a well-known ambiguity in the expression 'that which appears': if we say that 'x appears as y', the expression can either refer to x or y. Bird — as an example of deflationist interpreter who tries to explain away those passages where Kant seems committed to things in themselves — relies on this ambiguity and suggests that all Kant is implying here is that «[t]here are appearances, that is, things that appear» (Bird 2006, p. 560)[162]. Bird holds that we should understand the expression 'that which appears' in the Absurdity Claim as referring to y, i.e. the empirical object of intuition

---

[161] It is easy to count the passage (as others quoted below) as a textual support in favor of dual-aspect readings of transcendental idealism. At the same time, though, Kant's reference to objects of experience and objects as things in themselves as «*eben dieselben*» is potentially misleading for reasons I have already discussed in section 2.4. I thus prefer here to abstract as much as possible from the debate between 'dual-aspect' and 'two-objects' readings, as I am indeed suspicious towards the very meaningfulness of this distinction, which, in their most plausibly qualified versions tend to come very close to each other (see note 2 of my Introduction).

[162] Bird's reconstruction is precisely aimed at undermining Adickes' reading according to which Kant would here be arguing for the actuality of things in themselves.

which we call appearance, not to *x*, i.e. a thing in itself[163]. Bird's reading reduces the Absurdity Claim to a trivial identity between the term 'appearance' and the expression 'something that appears', a conclusion that does not seem very plausible, especially if we consider that similar claims can be found in other, less ambiguous passages (that I discuss below). I will instead accept the more straightforward equation of 'something that appears' with *x*, which I read as something belonging to the non-sensible level of reality.

At the same time, though, attention should be brought to Kant's emphasis on the verb '***denken***', to which he attaches an important footnote[164]. Precisely that footnote is where Kant presents one of the most explicit formulations of his distinction between logical and real possibility. Kant's point in the footnote is that I do not need to prove the real possibility of a thing or — which are equated in the text — the objective validity of its concept in order to think of that thing. For us to merely think something, the less demanding logical possibility of its concept is enough[165]. Although he concludes the footnote by reminding the reader that we can establish the real possibility of something also by drawing from practical sources of cognition, what matters for us here is the following: from a strictly theoretical perspective, the BXXVI-VII passage shows no commitment to anything more than the *logical possibility* of objects as things in themselves (or of a non-sensible level of

---

[163] Bird proceeds then to show that, according to Kant, the notions of 'appearance' and of 'thing in itself' «are necessarily connected sides of the same coin, and both are dependent on the concept of an "object in general"» (Bird 2006, p. 561) — something which Kant himself writes at the end of the same page (BXXVII). On Bird's account, then, the passage proves that the thing in itself must be thought only because it is necessarily linked to the notion of 'appearance', for both are included in the generic conception of an object.

[164] Which I mentioned earlier in note 16.

[165] In light of the clarifications made in section 1.5, to 'think something' is best understood here as 'purporting to think something'. That is, 'thinking something' in this sense does *not* imply successful reference. Kant's point in the footnote is thus that for us to 'think something' in this sense it is sufficient that the purported object of thought is logically possible, i.e. that its concept is not self-contradictory.

reality). As we will see, other passages suggest a stronger commitment, but this is what emerges from the quoted passage[166].

Now a second question: What does the Absurdity Claim precisely entail? What is it supposed to prove? At a closer inspection, the argument of BXXVI-VII does not prove the existence of things in themselves. On the contrary, it starts from the assumption, not of their existence, but only of their logical possibility in order to suggest no more than that thinking a non-sensible level of reality is required if we are to apply the concept <appearance> correctly to empirical objects in general. In its formulation at BXXVI-VII, at least, the Absurdity Claim expresses first of all a sort of linguistic or merely conceptual constraint. The concept of <appearance> (y) entails a reference to an x that appears as y. Analogously, the word 'mother', for example, entails a conceptual connection to 'child'. However, this conceptual link per se proves nothing about the objective reality: by arbitrarily applying the concept <mother> to a woman we do not thereby prove the existence of a child. Analogously, by arbitrarily calling something an <appearance> we do not yet provide a metaphysical demonstration of the actuality or real possibility of 'something that appears'. The process works rather *in the opposite direction*: in order to establish that my employing the concept <appearance> is in fact meaningful — in the case at issue: in relation to empirical objects in general

[166] I thus suggest that, given the difficulties just discussed, we should refrain from counting the BXXVI-VII passage as an example of Kant's explicit commitment to the existence of things in themselves, as interpreters often do (e.g. Van Cleve 1999, p. 135, Allais 2015, pp. 69-70, and Stratmann 2018, p. 16).

— I need to prove in advance that a *real*, objective relation (of 'appearing', as it were) occurs[167].

But then the question of the utmost importance becomes: *Why should we call empirical objects 'appearances' in the first place?* The moment Kant decides to apply the concept <appearance> (*y*) to empirical objects he is already assuming that there is something that appears (*x*); he is assuming the objective reality of a metaphysical relation of '*erscheinen*' involving empirical objects. Yet at BXXVI-VII Kant provides no theoretical justification for such assumption. All this means that the passage cannot be taken to provide a successful demonstration of the existence of thing in themselves[168].

One last remark on the passage. As I have stressed, Kant does not grant there more than the logical possibility of objects as things in themselves. Note however that, according to what said above, if only logical possibility (or eventually LP+) was granted to non-sensible reality, then the subsumption of the concept of <empirical object> under that of <appearance> would amount as well to no more than a logical possibility (or LP+). That is to say, on this scenario all that would be established is that it is logically consistent to call

---

[167] I thus disagree with Allais, who writes: «we are not making an unjustified existential commitment, because Kant seems to think that it is a conceptual truth that where there is something that appears in a certain way to us there is something that has a way it is in itself (BXXVI, A251-2). Similarly, he thinks that it is a conceptual truth that what is relational requires something non-relational (A284/B340) and he thinks that appearances are entirely relational. […] Anything we can know to be true as a matter of logic we can know to be true of everything that exists, so if something exists and logic commits us to thinking that there is a way things are in themselves, then we can know that there is a way things are in themselves […] Kant thinks that something's being logically possible is not enough to tell us that it is really possible, however, something's being logically impossible is enough to tell us that it is really impossible. And Kant thinks that logic tells us that relations require something non-relational and appearances require something that appears» (Allais 2015, pp. 69-70). I have already discussed in section 2.3 the alleged analytical link between relational (outer) and non-relational (inner) properties. What is at issue here, instead, is the connection between 'appearance' and 'something that appears', and my point thus far has been that such conceptual link is bound to remain sterile unless one is able to justify the employment of the term 'appearance'.

[168] Allison (2004, in particular ch. 3 section I) as well warns us against taking this and other passages (some of which I mention below) as providing a successful argument for the existence of things in themselves. However, Allison suggests that such passages turn out unproblematic once we accept his methodological reading of Kant's transcendental idealism. Unlike him, though, I am not persuaded that his reading can survive passages like the ones from the *Prolegomena* discussed below, where Kant's commitment to the existence of things in themselves is explicit — though, in my view, unwarranted.

empirical objects in general 'appearances' (y) in light of the fact that it is not contradictory to conceive of another level of reality (x) that appears (as y). If that was the case, for all we know then it might as well be (although we would not be able to tell) that nothing really exists independently of our faculties, and thus that empirical objects are not appearances after all. While I find this agnostic conclusion — i.e. to the effect we cannot really know whether empirical objects truly are appearances or not — philosophically attractive (for reasons I will discuss later), it is hardly representative of Kant's ultimate views on the topic. It seems in fact to be a central and ubiquitous claim of his transcendental idealism that empirical objects are, without a doubt, nothing but appearances. If this is so, then he must accept the *existence* of noumenal reality as equally certain.

Some version of the Absurdity Claim can be found in other passages. For instance, in the *Prolegomena* we read:

> In fact, if we view the objects of the senses as mere appearances, as is fitting [*wie billig*], then we thereby admit at the very same time that a thing in itself underlies them, although we are not acquainted with this thing as it may be constituted in itself, but only with its appearance, i.e., with the way in which our senses are affected by this unknown something**.** Therefore the understanding, just by the fact that it accepts appearances, also admits to the existence of things in themselves [*gesteht auch das Dasein von Dingen an sich selbst zu*], and to that extent we can say that the representation of such beings as underlie the appearances, hence of mere intelligible beings, is not merely permitted but also unavoidable [*unvermeidlich*] (*Prol*. 4: 314-5).

Here Kant explicitly writes that inasmuch as we accept appearances we must also admit the existence of '*Dinge an sich selbst*'. Kant's claim has a similar structure to the Absurdity Claim in the B-Preface, and brings with it the same problems. Only, this time his commitment to the *existence* of things in themselves is explicit. Given the conceptual constraint between 'appearance' and 'something that appears', «if we view the objects of the senses as mere appearances» we must therefore admit the existence of something that appears in them. Such «unknown something» are of course things in

themselves, «mere intelligible beings» which «underlie the appearances», as the passage explicitly states.

Again, though, this is not an argument that can miraculously lead us from the mere assessment of a conceptual constraint to the objective reality of noumena. Instead, the existence of noumena is already entailed in a crucial claim made at the outset of the passage, namely that viewing the objects of the senses as mere appearances is «fitting». If calling empirical objects 'appearances' is «fitting», then noumenal reality must exist[169] — as it will become clearer later on. Again, however, no justification is provided for why such employment of the word 'appearance' is appropriate.

In a couple of pages from the first edition of the Phenomena and Noumena (PN) chapter where Kant states again the Absurdity Claim[170], we find a passage where he sheds a little more clarity on the latter issue:

> Sensibility and its field, namely that of appearances, are themselves limited by the understanding, in that they do not pertain to things in themselves, but only to the way in which, *on account of our subjective constitution*, things appear to us. This was the result of the entire Transcendental Aesthetic (A251, italics mine).

---

[169] See Strawson's insightful considerations on the concept of a «corrected view» entailed in the opposition between «appearance» and «reality» (see Strawson 1966, in particular Part Four, section 5).

[170] A bigger section of the pages A249-53 would be worth quoting and analyzing. However, this would raise further series of difficulties from which I prefer to abstract here. To give an example, we find there one of the rare occasions where Kant discusses the notion 'transcendental object', which, for reasons I have already made clear (see note 66), I choose not to discuss at length in the present work. Another complication arises from the fact that Kant's claim that «the word "appearance" must already indicate a relation to […] an object independent of sensibility» (A252) is made in the context of a passage which begins with a slightly polemical tone: «But the cause on account of which, not yet satisfied through the substratum of sensibility, one must add *noumena* that only the pure understanding can think to the *phaenomena*, rests solely on this» (A252), namely on the Absurdity Claim. In fact, in the subsequent paragraph Kant is eager to stress that «the concept of a *noumenon*» signifies only «the thinking of something in general, in which I abstract from all form of sensible intuition», and that we cannot prove that this concept signifies «a true object». Hence, «although our thinking can abstract from that sensibility, the question still remains whether it is not a mere form of a concept and whether any object at all is left over after this separation» (A252-3). All this makes it more complex to assess how firm of a commitment to non-sensible reality Kant is willing to express in these pages.

It might seem that Kant provides us here with a satisfactory solution to the previously posed problem: empirical objects are appropriately called appearances because their sensible properties depend on our subjective constitution. As Kant readily reminds us, this — namely the ideality of space and time as forms of sensibility — was the major achievement of the Transcendental Aesthetic. Yet this answer does not bring us much further. For a lot hinges on how to understand the word '*subjective*' in this context. Given that space and time are empirically real, by calling space and time and thus sensibility 'subjective', or 'transcendentally ideal', Kant cannot take them to depend on the subjective constitution of the empirical, embodied, spatiotemporal subject. Rather, Kant is arguably committed to the alternative view according to which space and time are dependent on the subjective constitution of a transcendentally real subject — or whatever element of noumenal reality may correspond to what we conceive as a 'subject'. This of course is an aspect of the problem of affection in Kant, on which I will spend more words in the next section. What is relevant for the moment is that it is far from obvious that the assessment of the subjectivity of space and time, which Kant takes to justify the employment of the word 'appearance' with reference to empirical objects in general, proves the existence of noumenal reality. Indeed, it seems to be the other way around, since it is very likely that by deeming space and time as 'subjective' in the transcendental sense Kant presupposes already the existence of a transcendentally real, non-sensible level of reality.

Before considering these issues more thoroughly, let us first go back to the *Prolegomena*, and precisely to a passage located at the end of Kant's overview on the Transcendental Dialectic. For there another element of complexity is added to the picture, i.e. reason's quest for the unconditioned:

> The sensible world is nothing but a chain of appearances connected in accordance with universal laws, which therefore has no existence for itself; it truly is not the thing in itself, and therefore it necessarily refers [*bezieht*] to that which contains the ground [*Grund*] of those appearances, to beings that can be cognized not merely as appearances, but as things in themselves. Only in the cognition of the latter can reason hope to see its desire for completeness in the progression from the conditioned to its conditions satisfied for once.

> […]
> We should, then, think for ourselves an immaterial being, an intelligible world, and a highest of all beings (all noumena), because only in these things, as things in themselves, does reason find completion and satisfaction, which it can never hope to find in the derivation of the appearances from the homogeneous grounds [*Gründe*] of those appearances; and we should think such things for ourselves because the appearances actually do relate to something distinct from them (and so entirely heterogeneous), in that appearances always presuppose [*voraussetzen*] a thing in itself, and so provide notice of such a thing, whether or not it can be cognized more closely. Now since we can, however, never cognize these intelligible beings according to what they may be in themselves, i.e., determinately — though we must nonetheless assume such beings in relation to the sensible world, and connect them with it through reason […] (*Prol*. 4: 354-5).

In these passages we find two separate lines of reasoning. First is the familiar claim that «the sensible world», or «appearances», «necessarily refers to», or «presuppose» some non-sensible grounds, i.e. things in themselves. I have insisted that this is not a successful argument for inferring the existence of noumenal reality as that would instead require a proper justification for the employment of the word 'appearances' to empirical objects.

We find however also a second line of reasoning, which I take to be to some extent independent from the Absurdity Claim. Kant tells us that «[o]nly in the cognition of the latter [i.e. things in themselves] *can reason hope to see its desire for completeness in the progression from the conditioned to its conditions satisfied for once* [italics mine]». This cognition is of course not attainable, as non-sensible reality is «a wholly unknown (which will always remain so)» (*Prol*. 4: 354). Nevertheless, he insists that we should «*think* for ourselves an immaterial being, an intelligible world, and a highest of all beings (all noumena), because only in these things, as things in themselves, *does reason find completion and satisfaction* [italics mine]». As we have seen, it is pivotal for Kant's critical enterprise not only to protect the interests of pure practical reason from materialism, but also to secure the legitimacy of reason's never-ending metaphysical quest for the unconditioned. And for this scope, we must be at least able to think a realm of things in themselves beyond the domain of sensibility, or better, we must secure its LP+. It seems to me that

this, and not the Absurdity Claim, is one of Kant's most appealing arguments for at least taking into consideration an alleged noumenal reality, though not for its existence. Very briefly, the point is that reason's inevitable metaphysical need leads us to entertain thoughts about noumenal reality as the only domain to which some unconditioned conditions could ever belong; and, fortunately enough, critical philosophy has been able to prove that a non-sensible level of reality is logically possible and that its real possibility, though theoretically indemonstrable, cannot be disproved as well (= LP+)[171]. Now, one may wonder whether Kant would be able to argue for such LP+ without some commitment to the existence of noumenal reality, a complex question to which I cannot attempt to provide a satisfying argument. What matters the most for the present investigation is again that, in these passages, Kant is *not* providing an argument that moves from the assumption of reason's strive for unconditioned conditions to the indubitable existence of noumena.

Finally, consider another version of the Absurdity Claim from the PN chapter:

> if we call certain objects, as appearances, beings of sense (*phaenomena*), because we distinguish the way in which we intuit them from their constitution in itself, then it already follows from our concept that to these we as it were oppose, as objects thought merely through the understanding, either the very same conceived in accordance with the latter constitution, even though we do not intuit it in them, or else other possible things, which are not objects of our senses at all, and call these beings of understanding (*noumena*) (B306, amended translation[172]).

---

[171] I take the opportunity here to distance myself from Friebe's (2014) own brand of methodological dual-aspect view, insofar as he states that the appearance-aspect of objects is ontologically more fundamental than the in-itself aspect. Let us put aside Kant's eventual commitment to things in themselves as the 'something that appears' and accept instead no more than their LP+. Even in this case, a crucial feature of non-sensible reality that I have been stressing here is that it is introduced, though only as a LP+, as the domain where the unconditioned grounds demanded by our reason could find place. A logically possible+ noumenal realm is therefore at least purported to be a level of reality which is ontologically more fundamental than that of possible experience — a feature that is of course relevant in the practical context as well.

[172] For the reason why this translation has been amended, see note 63. See also my remarks in section 2.4 on the problems that arise if one takes the expression «*eben dieselbe*» too literally.

Kant's point seems to be the following: if we call certain objects «beings of sense (*phaenomena*)», we thereby oppose to them «beings of understanding (*noumena*)». Yet first of all these lines do not inform us explicitly about the import of Kant's commitment to noumena, i.e. whether they indubitably exist or amount to a mere logical possibility. In any case, though, once again it is clear enough from the first line that Kant starts from an assumption for which no argument is provided here, namely that we can appropriately deem beings of sense as 'appearances'. Let us then finally focus on such crucial assumption.

5.3 Concluding Considerations: Kant's Story about Affection

I have suggested that the existence of noumenal reality is in fact presupposed by Kant because it is already entailed in the assumption that empirical objects in general can appropriately be called 'appearances'. The latter assumption is connected to others, i.e. that space and time are *subjective* and that we are *finite* beings. At a closer look, though, whether those assumptions entail the existence of noumenal reality, as I suggested, or not, hinges on issues such as the following: To what level of reality do our faculties (e.g. understanding, sensibility and reason) belong? Do they belong to the phenomenal or noumenal (ground of the) self? From which standpoint can we establish the finitude of the cognitive subject? It seems to me that, unfortunately, Kant is never explicit on such pivotal issues.

Outside the practical domain, Kant is reluctant to talk explicitly about our faculties as noumenal powers of a noumenal being[173]. It is therefore tempting to read Kant's account of the origin of experience as a foremost phenomenal story, at least with regard to the causal relations (e.g., most crucially, affection) entailed therein. On such an immanent reading, the aim of transcendental philosophy would thus be to account for the conditions of objective validity of representations and principles, which nevertheless,

[173] In the practical domain, instead, Kant is clear that, for instance, we must ascribe freedom to ourselves as noumena (see e.g. *CPrR* 5: 6, 5: 42 and 5: 47).

belong to an empirical, embodied subject endowed with faculties[174]. Within such phenomenal framework, the generation of experience through affection may be construed (very roughly) as follows: through spatiotemporally informed empirical intuition I am aware of some sensible stimuli; 'after' the contributions of understanding and reason, though, I am able to question the previous process and judge that one or more individual objects located in space and time, by getting in contact with my sensory apparatus, are in fact causally responsible for what must be described as the initial affection of my sensibility. On this account of affection, all the actors of the story belong to the phenomenal domain: the embodied cognitive subject endowed with sensibility and understanding, the affecting objects, and of course the relation of affection itself. Kant appears to have in mind something of the like when he formulates the Anticipation of Perception (see section 2.4); or when, for instance, at the beginning of the Transcendental Aesthetic, he writes:

> The effect of an object on the capacity for representation, insofar as we are affected by it, is **sensation**. That intuition which is related to the object through sensation is called **empirical**. The undetermined object of an empirical intuition is called **appearance** (A20/B34).

However, it is hard to deny that such account requires the assumption of non-sensible reality after all[175]. As we know, space and time are only the forms of intuition, and as such are in some sense subjective. Hence, although in a second stage we are able to describe a process of affection involving two kinds of objects located in space and time, i.e. my embodied self and some phenomena, it must nevertheless be admitted that there is an initial affection

---

[174] This empirical picture is reflected in the interpretative framework through which, among others, Kitcher (1999) attempts to deal with the problem of affection in Kant.

[175] Karin de Boer (2014 and 2020) argues that providing a philosophical account of non-phenomenal affection lies outside the scope of the *Critique*. The misleading assumption that such an account must be found leads, on her view, to the opposite mistakes of commentators like Allison on the one hand, and Langton on the other. Now, I find it plausible to hold that Kant does not address explicitly (or at least not explicitly *enough*) the issue of noumenal affection; also, I am sympathetic to the claim to the effect that Kant seems most often to exclude that we as cognizers are in the position to judge about it. However, what I shall argue in the following pages is that, regardless of whether Kant addresses it or not, and of whether he is in the position to do it or not, the problem of non-sensible affection nonetheless lies at the basis of Kant's system.

whose ground is not in itself spatiotemporal. At least — it must be stressed — *inasmuch as we hold that space and time are no more than subjective forms*, for the latter claim implies that the forms of space and time come into play only as a consequence of sensibility being affected. The complete story of the origin of experience must then start from the assumption that some 'section' of noumenal reality — that we may attempt to depict as a noumenal self endowed with some faculties[176], including spatiotemporal sensibility — is affected by something equally non-sensible[177].

The initial affection through which matter is spatiotemporally arranged, so to speak, cannot occur at the empirical level. Unlike what Kant calls 'empirical idealism', his transcendental idealism does *not* hold that space and time are merely subjective features of the *empirical self's* receptive faculty. For him space and time are empirically real, which means that, from an empirical standpoint, they are to be viewed as *objective* features of phenomena. Instead, Kant's claim that empirical objects are appropriately called appearances, or that space and time are *subjective* forms translates into the view that, at least from the transcendental standpoint, space and time are not objective features of reality, but rather only the *forms of a non-spatiotemporal receptive faculty for being affected by something which is equally non-spatiotemporal*.

In order to illustrate Kant's original assumption and its limits, let me bring to the fore an essential element of experience which such assumption is meant to account for. Experience presents some features which cannot be cognized a priori, namely *a posteriori* features. Even though there is a large space for conceptual and perceptual interpretation, there are also undeniable constraints in our mental life. In this precise instant, I just happen to have

---

[176] While for ease of exposition here and elsewhere I speak of 'noumenal self' and its faculties, it is debatable whether such expression is justified in the theoretical domain, for Kant. On this regard, keep in mind the considerations at the end of section 4.6 and the ulterior complication mentioned in note 154 .

[177] That appearances in general are the result of affection is suggested by Kant, for example, at B69, A44/B61 and *Prol*. 4: 314-5. Stang (2015) provides a plausible account of (double) affection in Kant. Stang's main focus, against the objections raised by Jacobi and others, is to argue that noumenal affection (noumena affecting the noumenal self) is compatible with the empirical one (empirical objects affecting the empirical self). As Stang admits, though, his defense is not meant solve all the complications involved in Kant's doctrine of affection. Indeed, while I am conceding here that Kant is committed to some kind of noumenal affection, I take it to be a problematic and partially unwarranted assumption.

certain perceptions and not others, and indeed I just happen to have perceptions at all instead of an ungraspable blank. Neither the fact that I am endowed with receptivity, nor the immanent logic of experience seem enough to explain the fact that some matter of experience is simply given (in such and such ways). How then to account for the givenness of a posteriori features of experience? How to account for the fact that experience, is accompanied by what may be called an element of perceived 'limitation' or 'passivity'?

Kant's way of dealing with this feature of experience is to posit a distinction between spontaneity and receptivity at the transcendental (meaning here: non-empirical)[178] level, thus allowing him to ground such a posteriori features in something truly independent from our faculties[179]. While my spontaneous faculties and the forms of intuition are responsible for the a priori features of reality, my receptivity indicates that I am in touch with a noumenal something which affects me. All we know about such affection is that the affected faculty imposes the form of space and time on its content. Empirical objects are just appearances because they are the effect of a transcendentally real, noumenal something on some transcendentally real, 'space-time informing', receptive faculties (plus, of course, the activity of understanding and reason). If empirical objects were entirely constituted through our faculties, we would not be able to account for the a posteriori features of experience — or so Kant must have thought[180]. If, on the other

---

[178] The term transcendental has different, though to some extent interconnected meanings. Here I am not employing it with reference to the a priori conditions of possible experience (cf. e.g. A56/B80-1). Rather, what I have in mind is Kant's usage of the term when he claims that something can be ideal or real from either the empirical or the transcendental standpoint: here the term 'transcendental' denotes the perspective achieved by abstracting from sensible conditions (cf. e.g. A28/B44). Similarly, Kant distinguishes between an empirical and a transcendental meaning of the expression «**outside us [*ausser uns*]**» (see A373).

[179] Here a further complications arises. What if (as Kant seems to allow at A359-60; cf. again note 154) matter and the self had the same noumenal ground? Would the existence of a single, self-affecting noumenon suffice for Kant's purpose of grounding the a posteriori features of experience? Is such possibility in general compatible with Kant's original assumptions? I leave the question unanswered.

[180] That the a posteriori features of experience are a crucial reason for why Kant is led to posit the existence of things in themselves has been brought up with special clarity by Piché (2004) and Rosefeldt (2022).

hand, empirical objects did not depend at all on our faculties, no a priori cognition would be possible: space and time must be the (transcendentally) subjective element of experience contributed to by a 'noumenal receptivity', for otherwise no a priori cognition would take place in the way Kant describes it.

Therefore, to sum up: classifying empirical objects as 'appearances', calling space and time 'subjective', distinguishing between spontaneity and receptivity, and presupposing the existence of noumenal reality are in Kant all *aspects of the same assumption*[181]. In this way Kant is able to solve the problem of *a priori* cognition while at the same time maintaining the ultimate irreducibility of the *a posteriori features* of reality to our faculties — a result which is desirable for Kant given his assumption that we are essentially finite beings that are unable to *produce* all of reality by ourselves[182], unlike an intuitive understanding would do.

Unlike one may suspect, the main problem with such set of presuppositions is not 'Trendelenburg's objection' to the ideality of space and time. Briefly put, the well-known objection states that Kant neglected to consider an alternative option to the mere subjectivity or mere objectivity of space and time. That is, by claiming that space and time cannot be features of things in themselves Kant would have thereby excluded without justification the third option on which space and time are forms of human intuition *as well as* objective properties of things in themselves. The objection relies on the

---

[181] This is not to say that Kant provides no argument at all for these features of his system. Consider the subjectivity of space and time. Kant has provided several reasons in the Transcendental Aesthetic for taking space and time to be mere forms of intuition. And yet what is ultimately unwarranted, I believe, is the full import of such conclusion and its precise place in Kant's account of the origin of experience. For instance, it is debatable whether the label 'subjective' would fit space and time if all there existed and could possibly exist were empirical objects to which, as Kant admits, space and time must truly, i.e. objectively be ascribed.

[182] Hence in the *Prolegomena*, to distinguish, against his critics, transcendental idealism from the idealism of Berkeley and Descartes, Kant renames it «formal» idealism (*Prol*. 4: 337 and 4: 375) in order to stress that only the forms of experience are ideal, while its material conditions are ultimately independent from us. Speaking of 'material' is tricky for the word 'matter' has different meanings in Kant. For instance, it could refer to sensation as the type of subjective mental representation which constitutes the 'matter' of experience, or to the physical concept of 'matter' whose features are discussed most at length in the *MFNS*. The meaning of 'matter' at issue here is equivalent to neither of those two, although a precise definition of what is at stake is a complex philosophical issue.

alleged fact that it is conceptually possible for some properties to pertain to a thing objectively (e.g. the redness of a rose) while at the same time depending on the perceiving subject. Trendelenburg's objection is answered by Rosefeldt (2013), who reverts its main thrust against itself, by claiming that the picture it suggests is not so intuitive after all, as it is in fact «absurd»[183]. For example, even from an empirical standpoint, it is erroneous and non-sensical to ascribe a visual property such as 'redness' to the rose per se; rather, it should be said that the «rose has the property of appearing red to subjects with our visual sensory system» (p. 759). In the same way, if Kant is correct that spatiotemporality is an intrinsically relational property in that it depends on our sensible constitution, then Trendelenburg's hypothesis that it could be at the same time a feature of reality in itself does not make much sense either. Now, while I would not go as far as to say that for things in themselves to be spatiotemporal is a logical contradiction, as Kant most likely believes, it is fair to concede that, far from presenting an intuitive picture, the neglected alternative lies on the edge of unintelligibility. Hence Kant's view can be accepted, if not fully, at least to the effect that the neglected alternative is not worth taking (or it is at best conceptually difficult for us to take) into serious consideration[184].

Having seen that this is not the problem at issue here, my main concern about Kant's original assumption can be resumed by the following points: (1) our concepts cannot successfully refer to the non-sensible domain; (2) none of the metaphysical principles available to us — in particular the PSR in the vest of the principle of causality — can be used to infer from general features of the domain of experience to the indubitable existence of noumena; (3) we can conceive of other logically possible scenarios to account for the a posteriori features of experience in a broadly Kantian framework.

Let us start from (3). One may object to my critical remarks that all that is entailed in Kant's general assumption of non-sensible reality is the claim that

---

[183] I submit that Rosefeldt's use of the idea of response-dependence to rebut the 'neglected alternative' objection can be accepted independently of his metaphysical-dual aspect reading of Kant's idealism (more extensively developed in Rosefeldt 2022).

[184] This kind of weaker version of Kant's denial of the spatiotemporality of things in themselves is discussed e.g. by Gardner (1999, pp. 64-73) and Jauernig (2021, pp. 140-1 and 242).

there is something aside from ourselves that is responsible for the matter of experience[185]. This is all in all an acceptable assumption, the objection may go; for what is the alternative? Here, I suggest, is a Fichtean-like alternative which I take to be as plausible as, and as indemonstrable as Kant's own solution. In my experience I oppose myself to an objective world. Also, my experience is pervaded by a sensation of limitation, of passivity: the world discloses to me as in some crucial ways independent of my will. In my investigations of empirical psychology or philosophy of mind I come to describe myself as an embodied self endowed with a receptive faculty through which I am affected by empirical objects existing independently of me. However, from the transcendental standpoint the activity of thinking of which I am (arguably) aware of is all which is real and ultimate. What I empirically perceive as a limitation in my activity, a passivity that leads me to posit the existence of empirically real objects affecting me from the outside, is in fact nothing but a lesser degree of such pure thinking activity[186]. On this alternative view, activity is transcendentally real; passivity, instead, is merely empirically real, but transcendentally ideal[187].

One may object that a 'self-limiting pure activity that generates all reality', or something of the like is a rather mysterious conception. I respond that the notion of intelligible being and non-sensible reality, although apparently more familiar, are nevertheless equally if not less intelligible, the main reason being that our notions of objectivity and of individuality are inevitably bounded to the structure of our discursivity. It may further be objected that such pure thinking activity ends up being a noumenon, for it is not a spatiotemporal being. This issue is trickier: one answer could be that such activity is no proper 'being' at all, but a condition for all 'being', and that the only feature it has in common with Kant's conception of noumenal reality is

---

[185] On this point, see Rosefeldt (2022, pp. 37-8) who maintains that admitting the existence of what he calls «response-independent properties» of objects commits us to the minimally informing claim that the source of a mental state is placed in an extra-mental object

[186] See Giovanelli (2011) for a study of the influence on post-Kantian philosophers of the idea present in Kant's Anticipation of Perception according to which reality is a matter of degree.

[187] In Leibniz, for instance, we find as well the idea that the monad, i.e. a simple substance, is only active since as its passivity is only ideal (see e.g. *New Essays on Human Understanding*, Book II, Chapter XXI, § 72, and *Monadology* §7).

that it is metaphysically more fundamental than empirical objects. In any case, what matters here is that on the alternative picture just sketched no affection would occur at the transcendental level. To be sure, it is not my point to argue for such a Fichtean-like view here, which I introduced only as LP+. My suggestion is merely that we are not forced to make the fundamental assumptions Kant makes to allow for both a priori and a posteriori features of experience. Indeed, I claim that the best attitude we should maintain on Kantian grounds towards the ultimate sources of experience and reality is *agnosticism*. Accordingly, we should remain agnostic on whether the difference between spontaneity and receptivity is transcendentally real, and whether calling empirical objects 'appearances' is ultimately appropriate[188]. From the theoretical standpoint, we should therefore limit ourselves to the claim that the aforementioned distinction and the related use of the word 'appearance' constitute a logically possible+ framework — and indeed an extremely useful one, inasmuch as this allows us to follow our metaphysical strive for further conditions, and, most importantly, to pursue our practical ends.

As for the other concerns I listed above, enough has been said on (1), namely the worry about the impossibility for our concepts (including the categories) to successfully refer to something outside of the field of possible experience. So let us conclude with a few remarks on (2). Given that, on Kant's view, the only synthetic a priori metaphysical principles that are valid cannot be employed outside of the sensible domain, on which grounds can he establish a necessary connection between the general features of empirical objects and some non-sensible level of reality? It might very well be the case that the a posteriori features of reality simply *do not have an explanation*, for a *metaphysical* explanation involving *real* grounds is something we are

[188] Another way to formulate my worries is by stressing that we are not entitled to make knowledge claims from a 'God's eye' view. Hence, although I agree with Allison (2004, in particular ch. 2) that much of Kant's efforts are directed against the «theocentric model of knowledge» which adopts such illegitimate perspective, I find it all the more striking that (unlike what Allison believes) Kant himself eventually adopts it when he posits an original receptivity and when he makes claims about the ultimate status of space and time.

warranted to expect only *within* the realm of possible experience[189]. It is hard to see on which grounds a metaphysical, second order inquiry on the non-sensible grounds of *the sensible realm as such* is legitimate at all. This, however, poses a related question: given the limits of the PSR, what justifies Kant's transcendental reflection on the grounds of a priori cognition? In defense of Kant, Piché (2004) argues that Kant is *as* justified in assuming the existence of some non-sensible reality *as* he is in committing himself to other transcendental notions such as, for instance, pure apperception: both are introduced in quality of explanantia for the possibility of experience. I leave it open here whether such comparison is appropriate, for an adequate answer would require a serious investigation of Kant's argumentative strategies throughout the *Critique*.

Something more can be said in Kant's partial defense against the worries I have raised. First, accusing Kant of overstepping the limits he has established runs the risk of turning into a self-defeating claim. For it is all but certain whether it is eventually possible to defend a Kantian account of the limitations of discursivity without endorsing the set of assumptions revolving around noumenal affection discussed thus far. Second, it could be invoked that all philosophical projects inevitably entail some basic assumptions on which to build, and neither Kant nor his critics are in a privileged position on this respect[190]. These considerations are meant to remind us that it is not that simple to retain what each of us may view as Kant's most valuable insights and exclude the rest. In order to move a piece of the puzzle, we need to rethink the whole picture, and surely not many are able to do better than Kant himself. His immediate successors, each with his own different perspective, have put an enormous effort on such a complex enterprise.

The modest aim of the present work has been only to provide some elements for an overall deflationary interpretation of Kant's transcendental

---

[189] As this passage from the *Groundwork* very effectively states: «we can explain nothing but what we can reduce to laws the object of which can be given in some possible experience. […] Now, where determination by laws of nature ceases, there all *explanation* ceases as well, and nothing is left but *defense*» (4: 459).

[190] See e.g. Jauernig (2021), section 5.10.2

idealism that attempts to be as coherent as possible, yet without refraining from taking note of some inevitable short-circuits.

Abbreviations of Kant's Works

A/B:	*Critique of Pure Reason*
Anth.:	*Anthropology from a Pragmatic Point of View*
BL:	*Blomberg Logic*
CPrR:	*Critique of Practical Reason*
DWL:	*Dohna-Wundlacken Logic*
Groundwork:	*Groundwork of the Metaphysics of Morals*
JL:	*Jäsche Logic*
MFNS:	*Metaphysical Foundations of Natural Science*
ML_1:	*Metaphysics* L_1
MD:	*Metaphysics Dohna*
MM:	*Metaphysics Mrongovius*
MK_3:	*Metaphysics Vigilantius*
OaD:	*On a Discovery Whereby Any New Critique of Pure Reason is to be Made Superfluous by an Older One*
OPA:	*The Only Possible Argument in Support of a Demonstration of the Existence of God*
Prol.:	*Prolegomena to Any Future Metaphysics That Will Be Able to Present Itself as a Science*
Refl.:	*Reflexion*
VL:	*Vienna Logic*
WRP:	*What Real Progress has Metaphysics Made in Germany since the Time of Leibniz and Wolff?*

Bibliography

Allais, L. (2015) *Manifest Reality. Kant's Idealism and his Realism,* Oxford: Oxford University Press.

Allison, H. E. (1973) *The Kant-Eberhard Controversy: An English Translation Together with Supplementary Materials and a Historical-Analytic Introduction of Immanuel Kant's on a Discovery According to which any New Critique of Pure Reason Has Been Made Superfluous*, Baltimore and London: Johns Hopkins University Press.

—— (2004) *Kant's Transcendental Idealism. An Interpretation and Defense; Revised and Enlarged Edition,* New Haven and London: Yale University Press.

Ameriks, K. (1992) 'The Critique of Metaphysics: Kant and Traditional Ontology', in P. Guyer (ed.), *The Cambridge Companion to Kant*, Cambridge: Cambridge University Press, 249-279.

—— (2000) *Kant's Theory of Mind: An Analysis of the Paralogisms of Pure Reason*, Oxford: Clarendon Press.

Anderson, R. L. (2017) 'Lucy Allais on Transcendental idealism', in *Philosophical Studies*, 174: 7, 1661-1674.

Bader, R. (2017) 'The Refutation of Idealism', in J. R. O'Shea (ed.), *Kant's Critique of Pure Reason. A Critical Guide*, Cambridge: Cambridge University Press, 205-222.

Betti, P. (2024) 'Kant's Mature Account of Monads as Objects in the Idea', in *The Southern Journal of Philosophy*, 1-17.

Bird, G. (2000) '"Kantian Humility: Our Ignorance of Things in Themselves". By Rae Langton. (Oxford: Clarendon Press, 1998. Pp. xiv + 232. Price £35.00)', in *The Philosophical Quarterly* 50: 198, 105-108.

—— (2006) *The Revolutionary Kant. A Commentary on the Critique of Pure Reason*, Chicago: Open Court.

Brandt Bolton, M. (2021) 'Kant's Amphiboly as Critique of Leibniz', in B. C. Look (ed.), *Leibniz and Kant*, Oxford: Oxford University Press, 211-32.

Chignell, A. (2007) 'Belief in Kant', in *The Philosophical Review*, 116: 3, 323-360.

—— (2010) ‘Real Repugnance and Belief about Things-in-Themselves: A Problem and Kant’s Three Solutions’, in B. J. Bruxvoort Lipscomb and J. K. Krueger, *Kant’s Moral Metaphysics. God, Freedom, and Immortality*, Berlin/Boston: De Gruyter, 177-209.

—— (2017) ‘Can’t Kant Cognize Himself? Or, a Problem for (Almost) Every Interpretation of the Refutation of Idealism’, in A. Gomes and A. Stephenson (eds.), *Kant and the Philosophy of Mind: Perception, Reason, and the Self*, Oxford University Press, 138-158.

Clewis, R. R. (2015) ‘Editor’s Introduction’, in R. R. Clewis (ed.), *Reading Kant’s Lectures*, Berlin/Boston: De Gruyter, 1-29.

Clusa, J. (2024) ‘On the Semantics of Kant’s Concept of Substance’, in *Revista de Estudios Kantianos*, 9: 1, 158-178.

de Boer, K. (2014) ‘Kant's Multi-Layered Conception of Things in Themselves, Transcendental Objects, and Monads’, in *Kant-Studien* 105: 2, 221-60.

—— (2020) *Kant’s Reform of Metaphysics: The Critique of Pure Reason Reconsidered*, Cambridge: Cambridge University Press.

Di Bella, S. (2013) ‘The Myth of the Complete Concept: Completeness and Individuation in Kant and Leibniz’, in S. Bacin, A. Ferrarin, C. La Rocca and M. Ruffing (eds.), *Kant und die Philosophie in weltbürgerlicher Absicht. Akten des XI. Internationalen Kant-Kongresses*, Berlin/Boston: De Gruyter, 309-322.

Dyck, C. (2014) *Kant and Rational Psychology*, Oxford: Oxford University Press.

—— (2017) ‘The Principles of Apperception’, in G. Motta and U. Thiel (eds.), *Immanuel Kant: Die Einheit des Bewusstseins (Kant-Studien Ergänzungshefte)*, Berlin/Boston: De Gruyter. 32-46.

Edwards, M. (2023) ‘Synthetic Attributes and the Schematized Categories’, in *Kantian Review*, 28: 1, 21-40.

Emundts, D. (2022) ‘How Does Kant Conceive of Self-Consciousness?’, in K. Schafer and N. F. Stang (eds), *The Sensible and Intelligible Worlds. New Essays on Kant’s Metaphysics and Epistemology*, Oxford: Oxford University Press, 158-184.

Förster, E. (2002) 'Die Bedeutung von §§ 76, 77 der "Kritik der Urteilskraft" für die Entwicklung der nachkantischen Philosophie [Teil 1]', in *Zeitschrift für philosophische Forschung*, 56: 2, 169-190.

Freitag, M. and Kraus, K. (2020) 'An expressivist interpretation of Kant's 'I think'', in *Noûs*, 1-23.

Friebe, C. (2014) 'Kants Transzendentaler Idealismus: Eine Verteidigung der „methodologischen" Zwei-Aspekte-Deutung', in *Allgemeine Zeitschrift für Philosophie*, 39: 1, 1-23.

—— (2022) 'Kant's Ontology of Appearances and the Synthetic Apriori', in *Kant-Studien*, 113: 3, 498-512.

Friedman, M. (2001) '*Matter and Motion in the Metaphysical Foundations and the First Critique: The Empirical Concept of Matter and the Categories*', in E. Watkins (ed.), *Kant and the Sciences*, 53-69.

—— (2013) *Kant's Construction of Nature: A Reading of the Metaphysical Foundations of Natural Science,* Cambridge: Cambridge University Press.

Gardner, S. (1999) *Routledge Philosophy GuideBook to Kant and the Critique of Pure Reason*, Routledge: London.

Giovanelli, M. (2011) *Reality and Negation - Kant's Principle of Anticipations of Perception*, Dordrecht: Springer.

Grier, M. (2001) *Kant's Doctrine of Transcendental Illusion*, Cambridge: Cambridge University Press.

Guenova, L. (2013) 'Leibniz, Kant, and the Doctrine of a Complete Concept', in S. Bacin, A. Ferrarin, C. La Rocca and M. Ruffing (eds.), *Kant und die Philosophie in weltbürgerlicher Absicht. Akten des XI. Internationalen Kant-Kongresses*, Berlin/Boston: De Gruyter, 335-346.

Haag, J. (2015) 'Faculties in Kant and German Idealism', in D. Perler (ed.), *The Faculties: A History*, Oxford: Oxford University Press, 198-246.

Heidemann, D. H. (2012) 'The 'I Think' Must Be Able To Accompany All My Representations. Unconscious Representations and Self-consciousness', in P. Giordanetti, R. Pozzo and M. Sgarbi (eds.), Kant', in *Kant's Philosophy of the Unconscious*, Berlin/Boston: De Gruyter, 37-59.

—— (2013) 'Dass Ich bin: Zu Kants Begriff des reinen Existenzbewusstseins', in S. Bacin, A. Ferrarin, C. La Rocca and M. Ruffing (eds.), *Kant und die Philosophie in weltbürgerlicher Absicht. Akten des XI. Internationalen Kant-Kongresses*, Berlin/Boston: De Gruyter, 153-164.

—— (2021) 'Kant's Supposed Realism about Things-in-Themselves', in B. Himmelmann and C. Serck-Hanssen (eds.), *The Court of Reason. Proceedings of the 13th International Kant Congress*, Berlin/Boston: De Gruyter, 515-523

Hogan, D. (2022) 'Schopenhauer's Transcendental Aesthetic', in K. Schafer and N. F. Stang (eds.), *The Sensible and Intelligible Worlds. New Essays on Kant's Metaphysics and Epistemology*, Oxford: Oxford University Press, 45-69.

Howard, S. (2022) 'Kant on Limits, Boundaries, and the Positive Function of Ideas', in *European Journal of Philosophy*, 30: 1, 64-78.

Jauernig, A. (2021) *The World According to Kant. Appearances and Things in Themselves in Critical Idealism*, Oxford: Oxford University Press.

Kitcher, P. (1999) 'Kant's Epistemological Problem and its Coherent Solution', in *Philosophical Perspectives*, 13, 415-441.

—— (2011) *Kant's Thinker*, Oxford: Oxford University Press.

Kohl, M. (2015) 'Kant on the Inapplicability of the Categories to Things in Themselves', in *British Journal for the History of Philosophy*, 23: 1, 90-114.

—— (2020) 'Kant on Cognizing Oneself as a Spontaneous Cognizer', in *Canadian Journal of Philosophy* 50: 3, 395-412.

—— (2023) *Kant on Freedom and Rational Agency*, Oxford: Oxford University Press.

Kraus, K. (2019) 'The Parity and Disparity between Inner and Outer Experience in Kant', in *Kantian Review* 24: 2, 171-195.

—— (2020) *Kant on Self-Knowledge and Self-Formation*, Cambridge: Cambridge University Press.

—— (2021) 'Kant's Argument Against Psychological Materialism in the Prolegomena', in P. Thielke (ed.), *Kant's Prolegomena: A Critical Guide*, Cambridge: Cambridge University Press.

Kreines, J. (2016) 'Things in Themselves and Metaphysical Grounding: On Allais' Manifest Reality', in *European Journal of Philosophy*, 24: 1, 253-266.

Langton, R. (1998) *Kantian Humility: Our Ignorance of Things in Themselves,* Oxford: Clarendon Press.

La Rocca, C. (1994) 'Introduzione', in I. Kant, *Contro Eberhard. La polemica sulla Critica della Ragion Pura*, trans. and ed. with an introduction by C. La Rocca, Agnano Pisano and Pisa: Giardini Editori e Stampatori in Pisa, 1-54.

—— (1999) *Esistenza e giudizio. Linguaggio e ontologia in Kant*, Pisa: ETS.

Longuenesse, B. (1998) *Kant and the Capacity to Judge: Sensibility and Discursivity in the Transcendental Analytic of the Critique of Pure Reason*, trans. by C. T. Wolfe, Princeton: Princeton University Press.

—— (2005) *Kant on the Human Standpoint*, Cambridge: Cambridge University Press.

—— (2017) *I, Me, Mine: Back to Kant, and Back Again*, Oxford: Oxford University Press.

Lu-Adler, H. (2021) 'Kant and the Principle of Sufficient Reason', in *Review of Metaphysics*, 74: 3, 301-330.

Marshall, C. (2010) 'Kant's Metaphysics of the Self' in *Philosophers' Imprint*, 10: 8, 1-21.

—— (2018), 'Never Mind the Intuitive Intellect: Applying Kant's Categories to Noumena' in *Kantian Review*, 23: 1, 27-40.

Massimi, M. (2017) 'What is this Thing Called 'Scientific Knowledge'? – Kant on Imaginary Standpoints And the Regulative Role of Reason', in *Kant Yearbook*, 9: 1, 63-64.

——(2022) 'Absolute Space as a Necessary Idea Reading Kant's Phenomenology through Perspectival Lenses', in M. McNulty (ed.), *Critical Guide to Kant's Metaphysical Foundations of Natural Science*, Cambridge: Cambridge University Press.

McLear, C. (2018) 'Motion and The Affection Argument', in *Synthese*, 195: 11, 4979-4995.

—— (2020a) '"I Am the Original of All Objects": Apperception and the Substantial Subject', in *Philosopher's Imprint*, 20: 26, 1-38.

—— (2020b) 'Kantian Conceptualism/Nonconceptualism', in E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*.

Messina, J. (2021) 'The Content of Kant's Pure Category of Substance and Its Use on Phenomena and Noumena', in *Philosopher's Imprint*, 21: 29, 1-22.

Oberst, M. (2017) 'Kant über Substanzen in der Erscheinung', Kant Studien, 108: 1, 1-18.

—— (2018) 'Analytische Erkenntnis der Dinge an sich: das Beispiel der einfachen Teile', in V. L. Waibel, M. Ruffing and D. Wagner (eds.), *Natur und Freiheit. Akten des XII. Internationalen Kant-Kongresses*, Berlin/Boston: De Gruyter, 1331-1339.

O'Farrell, F. (1979) 'Intuition in Kant's theory of knowledge: Part 2: Intellectual intuition in Kant's Theory', in *Gregorianum*, 60: 4, 725-746

Onof, C. (2011) 'Thinking the In-itself and Its Relation to Appearances', in D. Schulting and J. Verbugt (eds.), *Kant's Idealism. New Interpretations of a Controversial Doctrine*, Dordrecht: Springer, 211-235.

Ottaviani, O. (2018) 'Noumenorum Non Datur Scientia. Kant and the Notion of Intelligible World: Between Monadology and Platonism', in *Con-textos Kantianos* 1: 7, 427-57.

Piché, C. (2004) 'Kant and the Problem of Affection', in *Symposium*: Canadian Journal of Continental Philosophy, 8: 2, 275-297.

Pollok, K. (2001) *Kants Metaphysische Anfangsgründe der Naturwissenschaft: ein Kritischer Kommentar*, Hamburg: Felix Meiner Verlag.

Proops, I. (2021) *The Fiery Test of Critique. A Reading of Kant's Dialectic*, Oxford: Oxford University Press.

Quarfood, M. (2004) 'The Thing in Itself. Methodological Perspective or Metaphysical Entity?', in M. Quarfood, *Transcendental Idealism and the Organism: Essays on Kant*, Stockholm: Almqvist & Wiksell, 16-65.

—— (2011) 'Discursivity and Transcendental Idealism', in D. Schulting and J. Verburgt (eds), *Kant's Idealism: New Interpretations of a Controversial Doctrine*, Dordrecht: Springer, 143-158.

Roche, A. (2010) Kant's Principle of Sense, in *British Journal for the History of Philosophy*, 18: 4, 663-691.

Rosefeldt, T. (2006) 'Kants Ich als Gegenstand', in *Deutsche Zeitschrift für Philosophie* 54: 2, 277-293.

—— (2013) 'Subject-Dependence and Trendelenburg's Gap', in S. Bacin, A. Ferrarin, C. La Rocca and M. Ruffing (eds.), *Kant und die Philosophie in weltbürgerlicher Absicht. Akten des XI. Internationalen Kant-Kongresses*, Berlin/Boston: De Gruyter, 755-754.

—— (2017) 'Subjects of Kant's First Paralogism', in A. Stephenson and A. Gomes (eds), *Kant and the Philosophy of Mind: Perception, Reason, and the Self*, Oxford: Oxford University Press, 221-44.

—— (2022) 'Being Realistic about Kant's Idealism', in K. Schafer and N. F. Stang (eds.), *The Sensible and Intelligible Worlds. New Essays on Kant's Metaphysics and Epistemology*, Oxford: Oxford University Press, 16-44.

—— (2024) ‘‘In Itself’: A New Investigation of Kant’s Adverbial Wording of Transcendental Idealism’, in *Kantian Review*, 1-19.

Schafer, K. (2022) ‘Kant's Conception of Cognition and our Knowledge of Things-in-Themselves’, in K. Schafer and N. F. Stang (eds.), *The Sensible and Intelligible Worlds. New Essays on Kant’s Metaphysics and Epistemology*, Oxford: Oxford University Press, 248-278.

—— (2023) *Kant’s Reason: The Unity of Reason and the Limits of Comprehension in Kant*, Oxford: Oxford University Press.

Schafer, K. and Stang, N. F. (2022) ‘Introduction’, in K. Schafer and N. F. Stang (eds.), *The Sensible and Intelligible Worlds. New Essays on Kant’s Metaphysics and Epistemology*, Oxford: Oxford University Press, 1-15.

Schulting, D. (2011) ‘Kant’s Idealism: The Current Debate’, in D. Schulting and J. Verbugt (eds.), *Kant’s Idealism. New Interpretations of a Controversial Doctrine*, Dordrecht: Springer, 1-25.

—— (2012) ‘Non-Apperceptive Consciousness’, in P. Giordanetti, R. Pozzo and M. Sgarbi (eds.), in *Kant’s Philosophy of the Unconscious*, Berlin/Boston: De Gruyter, 271-303.

—— (2017a) *Kant’s Radical Subjectivism: Perspectives on the Transcendental Deduction*, London: Palgrave Macmillan.

—— (2017b) ‘Kant's Idealism and Phenomenalism. Critical Notice of Lucy Allais's “Manifest Reality. Kant's Idealism and his Realism”’, in *Studi Kantiani*, XXX, 191-202.

—— (2018) *Kant’s Deduction From Apperception. An Essay on the Transcendental Deduction of the Categories*, Berlin/Boston: De Gruyter.

Schulting, D. (ed.) (2016) *Kantian Nonconceptualism*, London: Palgrave Macmillan.

Spagnesi, L. (2024) ‘Kant on Phenomenal Substance’, in *British Journal for the History of Philosophy*, 1-24.

Stang, N. F. (2015) ‘Who’s Afraid of Double Affection?’, in *Philosopher’s Imprint*, 15: 18, 1-28.

—— (2016) *Kant’s Modal Metaphysics*, Oxford: Oxford University Press.

—— (2021) ‘Kant and the concept of an object’, in *European Journal of Philosophy*, 29: 2, 299–322.

—— (2022) ‘Thing and Object: Towards an Ecumenical Reading of Kant’s Idealism’, in K. Schafer and N. F. Stang (eds.), *The Sensible and Intelligible*

Worlds. New Essays on Kant's Metaphysics and Epistemology, Oxford: Oxford University Press, 296-336.

Stratmann, J. (2018) 'Kant, Grounding, and Things in Themselves'', in *Philosopher's Imprint*, 18: 7, 1-21.

—— (2022) 'The End of Explanation: Kant on the Unconditioned', in *Philosophy and Phenomenological Research*, 105: 3, 507-532.

Strawson, P. F. (1966) *The Bounds of Sense: an Essay on Kant's Critique of Pure Reason*, London: Methuen.

Sturm, T. (2001) 'Kant on Empirical Psychology. How Not to Investigate the Human Mind', in E. Watkins (ed.), *Kant and the Sciences*, Oxford: Oxford University Press, 163-184.

Tilliette, X. (1995) *L'intuition Intellectuelle de Kant a Hegel*, Paris: Vrin.

Timmermann, J. (2010) 'Reversal or Retreat? Kant's Deductions of Freedom and Morality', in A. Reath and J. Timmermann (eds.), *Kant's Critique of Practical Reason: A Critical Guide*, Cambridge: Cambridge University Press.

Van Cleve, J. (1999) *Problems from Kant*, Oxford: Oxford University Press.

Watkins, E. (1998) 'The Argumentative Structure of Kant's Metaphysical Foundations of Natural Science', in *Journal of the History of Philosophy*, 36: 4, 56-93.

Watkins, E. and Willaschek, M. (2020) 'Kant on Cognition and Knowledge', in *Synthese* 197: 8, 3195-3213.

Willaschek, M. (1998) '13. Phaenomena/Noumena und die Amphibolie der Reflexionsbegriffe: (A235/B294–A292/B349)', in G. Mohr and M. Willaschek (eds.), *Immanuel Kant: Kritik der reinen Vernunft*, Berlin: Akademie Verlag.

—— (2001) 'Die Mehrdeutigkeit der kantischen Unterscheidung zwischen Dingen an sich und Erscheinungen Zur Debatte um Zwei-Aspekte- und Zwei-Welten- Interpretationen des transzendentalen Idealismus', in R. Schumacher, R. P. Horstmann and V. Gerhardt (eds.), *Kant und Die Berliner Aufklärung: Akten des Ix. Internationalen Kant-Kongresses. Bd. I: Hauptvorträge. Bd. II: Sektionen I-V. Bd. III: Sektionen VI-X: Bd. Iv: Sektionen XI-XIV. Bd. V: Sektionen XV-XVIII*, Berlin/Boston: De Gruyter, 679-690.

—— (2018) *Kant on the Sources of Metaphysics: The Dialectic of Pure Reason*, Cambridge: Cambridge University Press.

Wolf, W. C. (2023) 'Kant's "in itself": Toward a New Adverbial Reading', in *Kant-Studien*, 114: 2, 207-246.

Wuerth, J. (2014) *Kant on Mind, Action, and Ethics*, Oxford: Oxford University Press.

Zanette de Araujo, R. (2023) 'Things in Themselves and the Inner/Outer Dichotomy in Kant's Amphiboly of the Concepts of Reflection', in *Kantian Review*, 28: 1, 143-156.

Zinkin, M. (2008) 'Kant's Argument in the Amphiboly', in V. Rohden, R. R. Terra, G. A. de Almeida and M. Ruffing (eds), *Recht und Frieden in der Philosophie Kants: Akten des X. Internationalen Kant-Kongresses*, Berlin: De Gruyter, 845-55.