

TRANSCENDENTAL IDEALISM  
AND THE ORGANISM



ACTA UNIVERSITATIS STOCKHOLMIENSIS  
STOCKHOLM STUDIES IN PHILOSOPHY 26

TRANSCENDENTAL IDEALISM  
AND THE ORGANISM

ESSAYS ON KANT

Marcel Quarfood

ALMQVIST & WIKSELL INTERNATIONAL  
STOCKHOLM

Doctoral Dissertation

Department of Philosophy, Stockholm University, S-106 91 Stockholm

## ABSTRACT

The notion of the organism has a somewhat ambiguous status in Kant's philosophy. On the one hand it belongs to natural science, on the other hand it is based on an analogy with the structure of reason. Biology therefore has a peculiar place among the sciences according to Kant: it is constituted by the use of a regulative maxim.

The present study places Kant's views on biological teleology in the larger context of transcendental idealism. It consists of five essays. The first one treats the notions of things in themselves and appearances, arguing for an interpretation in terms of two aspects or perspectives rather than two worlds. The importance of the discursivity of our cognitive capacity is stressed, as well as the need to separate Kant's various reflective perspectives. In the second essay this interpretation is applied to the third section of the *Groundwork*, arguing that this text does not belong to theoretical metaphysics, but rather to the articulation of a specifically practical perspective. The third essay discusses similarities and differences between Kant's a priori conditions for cognition and conceptions of innate ideas in the rationalist tradition. Kant's comparison of the system of categories with the biological theory of epigenesis is considered in connection to eighteenth century theories of generation. The comparison is viewed as an analogy rather than as a naturalistic theory of the a priori. In the fourth essay Kant's account of functional attribution in biology is explicated in the context of the present day debate of the issue. It is claimed that Kant's neo-Aristotelian approach avoids some of the difficulties in the dominant naturalistic accounts of today. Kant's view differs from the Aristotelian in that it involves a distinction of levels, making it possible to take functional attributions on the one hand as objective from the standpoint of biology but on the other hand as having a merely regulative status from a philosophical point of view. In the fifth essay an interpretation of the antinomy of teleological judgment in the *Critique of Judgment* is offered. The antinomy is taken to consist in the dialectical tendency to treat the regulative maxims of teleology and mechanism as constitutive principles. The difference between the discursivity of the human understanding and the idea of a non-discursive understanding, an important theme in Kant's solution of the antinomy, puts the question of biological teleology in relation to central tenets of transcendental idealism.

Key words: Kant, things in themselves, transcendental idealism, perspective, intelligible world, epigenesis, *acquisitio originaria*, biological functions, teleology, organism, antinomy, regulative maxim, discursivity.

© Marcel Quarfood 2004

ISBN 91-22-02099-3

Printed by Akademytryck AB, Edsbruk 2004

ISSN 0491-0877

## TABLE OF CONTENTS

Acknowledgments .....	7
INTRODUCTION .....	9
THE THING IN ITSELF: METHODOLOGICAL PERSPECTIVE OR METAPHYSICAL ENTITY? .....	16
1. Transcendental Idealism and the Thing in Itself: Two Worlds or Two Aspects? .....	17
2. Metaphysical and Methodological Views .....	34
3. Things in Themselves in Perspective .....	37
4. Critique of Metaphysical Interpretations .....	45
5. Concluding Remarks: Ethical Dimensions .....	61
KANT'S PRACTICAL DEDUCTION OF MORAL OBLIGATION IN <i>GROUNDWORK</i> III .....	66
1. Negative and Positive Freedom .....	67
2. Acting Under the Idea of Freedom .....	68
3. The Circle and the Two Standpoints .....	69
4. The Theoretical Dimension .....	73
5. Some Final Considerations .....	74
<i>ACQUISITIO ORIGINARIA</i> AND EPIGENESIS: METAPHORS FOR THE <i>A PRIORI</i> .....	77
1. Innatism and <i>Acquisitio Originaria</i> .....	80
2. The Epigenesis of Pure Reason .....	89
3. A Biological Interpretation of the <i>A Priori</i> .....	102
4. Concluding Remarks: Reason and the Organism .....	115
BIOLOGICAL FUNCTIONS IN A KANTIAN PERSPECTIVE.....	118
1. Etiological Functions .....	119
2. Causal Role Functions .....	128
3. Restricting Causal Role Functions to Biological Systems .....	133

*Table of contents*

4. The Normal, the Good, and Aristotelian Function .....	139
5. Kant's Notion of the Organism .....	145
6. Normative Function as Regulative Maxim .....	148
7. Conclusion: Kant after Darwin.....	157
THE ANTINOMY OF TELEOLOGICAL JUDGMENT .....	160
1. Presentation of the Antinomy .....	161
2. Programme for an Interpretation .....	166
3. The Dogmatic Systems: §§ 72–73.....	172
4. The Concept of Natural Purpose Is Not Objective: §§ 74–75... 175	
5. The Peculiarity of the Human Understanding: §§ 76–77 .....	177
6. Unification of the Principles: § 78 .....	191
7. Mechanism and the Inexplicability of Organisms .....	196
8. A Final Problem.....	205
WORKS CITED.....	209

## Acknowledgments

It has taken a surprisingly long time to complete the present thesis. Among all those to whom I am indebted for encouragement of various kinds while composing it, I wish in particular to mention the following. Dugald Murdoch, who as supervisor had the unenviable task of going through the fragmentary drafts of these essays, has given me good advice on philosophical matters as well as on questions of language and style. The final result has been much improved thanks to his efforts; needless to say, I am the one responsible for all remaining shortcomings. I am grateful to Gunnar Svensson, Chairman of the Department of Philosophy in Stockholm, and Dag Prawitz for making arrangements that made it possible for me to bring the work to a conclusion. I don't remember whether it was Gunnar or Dag who first proposed that I should present the thesis in the form of a set of essays rather than as a monograph. In any case, this turned out to be a good idea. My interest in Kant was originally awakened in the Kant seminars conducted by Alexander Orłowski, who also was my supervisor at the beginning of this project. With Markku Leppäkoski I have had many stimulating discussions on all things Kantian over the years. He has also commented on parts of the manuscript. Others who have read one or another part and given valuable advice are Helen Belcastro, Robert Callergård, Sören Häggqvist, Mikael Janvid, and Paul Needham.

In March 2000 I attended the IX International Kant Congress in Berlin, supported by a travel stipend from Wallenbergsstiftelsen. There I presented what is now the second essay of this thesis, "Kant's Practical Deduction of Moral Obligation in *Groundwork* III." Dieter Schönecker offered some good criticism on that occasion. The essay was published in the proceedings of the congress (*Kant und die Berliner Aufklärung*, vol. 3, ed. by V. Gerhardt, R.-P. Horstmann, and

*Acknowledgments*

R. Schumacher, Berlin: de Gruyter 2001), and is reprinted here with permission from Walter de Gruyter GmbH & Co.

Finally I thank my wife Clara Lindberg and our daughters Joanna and Merete, who have had to put up with much during this time.



# Introduction

The original aim of the present study was to offer a comprehensive exposition of the role of teleology in Kant's theoretical philosophy. There are two important themes for such an investigation, one local and one global. Locally, teleology has a place in Kant's account of natural science. First and foremost, biology seems to require (or at least to make free use of) a teleological notion of the organism, which Kant explicates in the concept of a "natural purpose" (*Naturzweck*). But also general methodological maxims (such as the principle of parsimony) are thematized in teleological terms, as aspects of a "purposiveness of nature." Now, this more general invocation of purposiveness leads over to the second, global theme: the place of teleology in the articulation of Kant's system of philosophy. In the *Critique of the Power of Judgment*, the purposiveness of nature is assigned to the reflective power of judgment. In distinction to the determinative power of judging, which produces judgments ascribing objective properties to things, reflection envisages an order in nature that can only be subjectively justified, but nonetheless is a necessary presupposition guiding all cognitive pursuits.<sup>1</sup> Since the distinction between the reflective and the determinative employment of the power of judgment is introduced only in the third *Critique*, exegetical questions arise concerning the relation of this distinction to Kant's less developed and rather different treatment of reflection in the first *Critique*, and in general as to its consistency with the division of the faculties. Such architectonic questions may well have large consequences for the global interpretation of transcendental philosophy. Another important teleological theme in Kant's architectonic is his

---

<sup>1</sup> *Critique of the Power of Judgment*, Ak. 5:185: "The power of judgment thus also has in itself an *a priori* principle for the possibility of nature, though only in a subjective respect, by means of which it prescribes a law, not to nature (as autonomy), but to itself (as heautonomy) for reflection on nature."

## Introduction

conception of philosophy as systematic. His notion of system is that of an organized body of cognition, in analogy with the functional build of an organism. Reason itself is conceptualized in such terms; given this prominent place for teleology, one must suspect that it should be of some importance for understanding transcendental idealism.<sup>2</sup>

To cover the above-mentioned issues in a monograph turned out to be a rather demanding enterprise, not least in view of the need to provide a general account of Kant's theoretical philosophy as a background. Moreover, considering that Kant appeals to teleology as a bridge between the theoretical and the practical domains,<sup>3</sup> the somewhat artificial limitation of the topic to theoretical philosophy presented further difficulties. It was therefore refreshing to abandon the original plan and set out to treat some of the issues in the format of a collection of self-contained essays. As it were, instead of a novel I ended up with a set of short stories. This change allowed the use of different approaches. Though I mostly develop an internally Kantian interpretation of various texts and themes, in one case an attempt is made to put Kant's ideas to work in the context of present day philosophy of biology.<sup>4</sup> However, it should be noted that my main intention is to reach a reasonable degree of coherence in the interpretation of the issues under study, guided by a general idea on how to approach Kant's philosophy. The task of criticizing Kant's thought has been left aside in favour of (more or less reconstructive) interpretation (which in any case is a necessary condition for the possibility of criticism).

As a consequence of the decision to let the essays be self-contained, some amount of repetition was unavoidable. My hope is that this is not entirely bad, as it allows various topics to be viewed in more than one context.

The general point of view on Kant's system adopted here is that it consists of co-existing domains which to a large extent are autonomous, and yet also entwined in complex ways. In particular, this

---

<sup>2</sup> It has for instance been argued that a central argument in the *Critique of Pure Reason*, the Transcendental Deduction of the categories, rests on teleological considerations. See Rosales 1989 and Dörflinger 1995.

<sup>3</sup> See for instance Düsing 1986 and Guyer 2001 on this theme.

<sup>4</sup> See the fourth essay, "Biological Functions in a Kantian Perspective."

applies to the distinction between the practical and the theoretical domains.<sup>5</sup> In addition to these fundamental standpoints or perspectives, one can find further salient perspectival distinctions, such as that between the empirical and the transcendental levels of investigation in theoretical philosophy. It is in no way original to stress the importance of questions of perspective in interpreting Kant. In fact, there is hardly one commentator who would neglect to pay some attention to such issues. Nevertheless, my impression is that the notion of distinct perspectives should be taken even more seriously in order to understand Kant better. However, I do not attempt to present a general account of the interconnection of perspectives in Kant's system.<sup>6</sup> Rather than becoming fully developed, the idea serves as a point of view from which to tackle other interpretive questions.

It is now time to give a short preview of the content of the essays. The first one, "The Thing in Itself: Methodological Perspective or Metaphysical Entity," defends the two-aspect view on the distinction between things in themselves and appearances. This view has recently come under attack, and it is now widely felt that the two-aspect view, at least when construed as an ontologically neutral epistemological distinction, is inadequate for the understanding of metaphysical claims implicit in Kant's practical philosophy, as well as for addressing other legitimate ontological questions. I disagree with this verdict on both textual and systematic grounds, and attempt to present Kant's distinction in terms of different epistemic perspectives, where one is the perspective we have as finite beings endowed with receptivity and spontaneity, whereas the other one is the idea of an entirely spontaneous cognition that does not need to receive data from outside. The idea of such an "intellectual intuition" serves as a contrast to highlight the character of the human "discursive" understanding. The essay can be seen as a defence of what I take to be the gist of the two-aspect view in the version of one of its originators, Henry Allison.<sup>7</sup> The defence includes criticism of

---

<sup>5</sup> There are also related ideas, such as that of different "interests of reason." See for instance *Critique of Pure Reason* A466/B494 and A666/B694.

<sup>6</sup> For one such attempt (which sometimes goes too far in this direction even for my taste), see Palmquist 1993.

<sup>7</sup> See Allison 1983.

some recent examples of metaphysical interpretations.<sup>8</sup> One thought guiding my treatment of the issue is that one should not accept an assumption that proponents of metaphysical views on the thing in itself tend to make, namely that standard metaphysical questions and concepts are neutrally applicable to any philosophical system. It seems to me that part of Kant's Copernican revolution consists precisely in the rejection of such neutrality, in favour of the critical approach which focuses on conditions for epistemic access rather than on ontology. Now, Kant may very well be wrong as to the possibility of such a revolution. What seems implausible is only the notion that the critical philosophy should be interpreted along metaphysical lines taken to be independently valid, as if this philosophy did not involve a sustained attempt to challenge the supposed neutrality of metaphysics.<sup>9</sup>

The first essay ends with a short discussion of the practical employment of the notion of things in themselves, and this theme is continued in the second essay, "Kant's Practical Deduction of Moral Obligation in *Groundwork* III." This foray into practical philosophy attempts to show that even the third section of the *Groundwork*, in which Kant appeals to the "intelligible world" in terms that apparently transcend all critical restrictions, can be read on the lines of the two-aspect interpretation. The key to this reading is the idea of a distinctive practical perspective, in which we "act under the idea of freedom."<sup>10</sup> The essay, originally written for and presented at the IX International Kant Congress in Berlin in 2000, is short and a bit programmatic. It indicates a possible way of reading Kant's text but

---

<sup>8</sup> In particular the one given in Langton 1998.

<sup>9</sup> I must here add the proviso that on certain interpretations of "metaphysical," it may very well be said that Kant's philosophy is itself metaphysical. It is for instance plausible to hold that transcendental philosophy was intended to replace rationalist metaphysics, and in that sense amounts to a new metaphysics. Also, it can be argued that metaphysical decisions are unavoidable on a fundamental metacritical level, when it comes to determining the properties of the cognitive capacities assumed as starting point for the critical undertaking. What is less plausible is the assumption that rationalist or modern metaphysical tenets should be frictionlessly applicable to Kantian philosophy, without themselves being put in jeopardy in the process.

<sup>10</sup> *Groundwork for the Metaphysics of Morals*, Ak. 4:448.

does not offer a fully worked out interpretation of the many complex issues raised in *Groundwork* III.

The first two pieces thus touch upon fundamental questions on the nature of transcendental idealism and the relation between theoretical and practical applications of reason. The third essay, “*Acquisitio originaria* and Epigenesis: Metaphors for the *A Priori*,” treats another central concept in Kant, that of the *a priori*. I try to show that Kant’s *a priori* is not clarified by its assimilation to rationalist conceptions of innate ideas, even though certain points of contact with the Leibnizian view on innateness can be found. Two metaphors deployed by Kant in explaining his view are examined, that of *acquisitio originaria* (a concept from the sphere of jurisprudence) and that of *epigenesis* (a biological concept of the organism’s gradual self-development). As Kant also took an interest in the biological questions raised by the concept of epigenesis, I look into some of his biological views and compare them to positions in the eighteenth century debate on the theory of generation. One aim here is to show that the appeal to epigenesis cannot be taken to imply that transcendental philosophy is grounded in biology. The essay ends with a brief discussion of the signification of the analogy between reason and the organism.

The third essay leads over from a fundamental question of transcendental idealism (the nature of the *a priori*) to local questions in Kant’s philosophy of biology. The fourth essay continues the exploration of these latter questions, but this time in a modern context. In a survey of developments in the recent debate on biological functions, I argue that the most favoured alternative on offer, the etiological theory, which bases functions exclusively on evolutionary adaptations, does not quite fulfil its goal of naturalizing teleological attributions in biology. As shown by Paul Sheldon Davies,<sup>11</sup> it moreover seems that the etiological concept of function is reducible to its main rival, the causal role account. This account takes an organizational view on functions, considering them to be causal contributions of component systems to a larger encompassing system. What seems to be lacking in this account, however, is an explanation of why particular systems (namely the biological ones) are viewed in teleological terms, whereas other systems are not. A third type of account, the neo-Aristotelian, has the potential to overcome the

---

<sup>11</sup> Davies 2000b.

## Introduction

limitations of the aforementioned theories, because it considers it legitimate to appeal to the idea of an organism's having a good, so that functions can be defined as what contributes to attaining this good. Though the neo-Aristotelian account has this advantage over its competitors, it has difficulties on the side of ontology. It seems to postulate objective biological values which are problematic from the naturalistic point of view usually adopted in the debate. It is at this point that Kant's views on teleological attribution can contribute to the discussion. As I reconstruct the teleological concept of natural purpose (*Naturzweck*) developed in the *Critique of the Power of Judgment*, it can be equated with the Aristotelian conception. The difference between Kant and Aristotle lies on the level of ontology: Kant takes an agnostic stance on the objectivity of the biological values postulated in the Aristotelian account. In this he is closer to modern naturalism, which prefers to avoid postulating values. But Kant differs also from the naturalist, in that the latter denies the existence of values rather than being agnostic about them. Kant's position, invoking the distinction between the reflective and the determinative power of judgment, is that biological judging is of the reflective kind, making use of a teleological (value-ascribing) point of view which is regulative but lacks definite ontological commitment. The teleological perspective serves to identify organisms, and thereby to constitute a particular field of study, namely biology. But if teleology is constitutive of the possibility of biology, how can this be reconciled with the claim that its status is merely regulative? My suggestion is that Kant's position should be taken to involve a perspectival distinction. On the "object level," Aristotelian teleology is ascribed to organisms and their functional parts. The object level, the biological field of study, is constituted by this use of teleology. On the other hand, on the "meta level," the level of philosophical reflection, ontological commitment to biological values is avoided and teleology is accorded merely regulative status.

The fifth and last essay, "The Antinomy of Teleological Judgment," consists of a reading of the *Dialectic of the Teleological Power of Judgment*, a part of Kant's third *Critique* which has proved difficult to interpret satisfactorily. In a way, its main ideas are quite simple. An antinomy arises between the mechanistic principle on the one hand and the principle that certain natural objects (that is, organisms) should be considered from a teleological point of view on

## *Introduction*

the other hand. The solution consists in dropping all ontological assumptions and putting the principles to use as mere regulative maxims. As shown in an excellent study by Peter McLaughlin,<sup>12</sup> this reading of the antinomy is a bit too easy. It cannot explain the complex structure of the text, and there remain questions concerning the possibility of opposed regulative maxims to co-exist. Moreover, a conflict threatens to arise between the claim of the first *Critique* concerning the constitutive status of the causal principle and the alleged regulative status of mechanism. My interpretation attempts to defend a version of the “easy” solution outlined above, while paying regard to the many insights of McLaughlin’s analysis. The problem of the co-existence of conflicting maxims is considered from the point of view of the object and meta level distinction sketched in the foregoing essay.

In his exposition of the solution to the antinomy, Kant inserts two unusually complex sections, §§ 76–77, in which reflection on the special character of the human discursive cognitive capacity is brought to bear on the case of teleological judgment. This puts the question of teleology in the larger framework of transcendental philosophy, and the investigation of these sections offers an occasion to return to considerations touched upon in the first essay concerning the contrast between discursive and non-discursive intellects and the function of this contrast in the articulation of critical philosophy.

---

<sup>12</sup> McLaughlin 1990.

# The Thing in Itself: Methodological Perspective or Metaphysical Entity?

There is no shortage of commentaries on Kant's concept of the thing in itself. Some would say there are too many. That I nevertheless venture to broach the subject is because I think the strength of the so called two-aspect interpretation is somewhat underestimated. Not so long ago this way of understanding the issues that surround the notion of things in themselves changed many philosophers' perception of Kant's philosophy. In my view, the picture that emerged was that of a more interesting thinker, less stuck in implausible speculations about the metaphysical properties of the thing in itself, and more relevant to developments in contemporary philosophy, analytic as well as continental.

However, in the last ten years or so the trend among Kant specialists (though other philosophers may not yet be aware of it) is to revert to a more metaphysical approach. Since the new interpretations are often very sophisticated in terms of exegetical scholarship, it is not easy to adjudicate between this trend and the two-aspect approach. Both can find textual support. Moreover, these metaphysically oriented readings often construe Kant's distinction between thing in itself and appearance in terms of two aspects, so that the contrast to the standard two-aspect view must be redrawn. To distinguish the two kinds of two-aspect approach, they are labelled the "metaphysical" and the "methodological" view, respectively, and the perspectival nature of the latter is stressed.

What is suggested in this essay is that the resources of the methodological two-aspect view are currently underrated, and that the motivations for a robustly metaphysical reading of the notion of the



thing in itself, to the extent that they rest on legitimate desiderata, can be met also on this interpretation.

The structure is as follows. In the first section, the two-world and the two-aspect approaches are presented, with a brief outline of transcendental idealism offered as a background to the latter view. Section 2 deepens the discussion by introducing the recent framing of the debate in terms of a distinction between two kinds of two-aspect view, the metaphysical and the methodological interpretation, of which the first is seen as a more sophisticated descendant of the two-world interpretation. In the third section, two versions of the methodological view are presented which rely on the notion of perspective as a crucial tool for the interpretation of Kant. Section 4 deals critically with some recent examples of metaphysical interpretations. The last section rounds off the discussion with a quick glance at the use of the concept of the thing in itself in Kant's practical philosophy.

## 1. TRANSCENDENTAL IDEALISM AND THE THING IN ITSELF: TWO WORLDS OR TWO ASPECTS?

Discussions of Kant's philosophy tend to be less fruitful if the question of the thing in itself and its paradoxical status (as unknowable and still, somehow, known to exist) is taken as a matter that can be discussed apart from Kant's larger claims about transcendental idealism. The ideal approach would probably be to consider the various statements about things in themselves as conclusions rather than premises, and to show in detail how they are reached in their argumentative contexts. I will not be able to accomplish such a demanding task, but at least some of these contexts will be presented. First, something should be said about the doctrine of transcendental idealism.

The nature of Kant's idealism has always been a contested issue, and interpretations differ greatly. What does he mean by qualifying his idealism as "transcendental"? The term "transcendental" is defined as what concerns not objects directly, but rather our way of cognizing objects *a priori* (A12/B25). Transcendental idealism would

then be an idealism with regard to our *a priori* mode of cognition. This mode of cognition has to do largely with the interplay between space, time, and the categories, these being factors which Kant traces back to the cognitive constitution we have as beings endowed with a passive (receptive) sensibility and an active thinking. So what Kant says is that the ideality of his idealism pertains to these cognitive conditions, and not to objects *per se*. In the *Prolegomena* (1783), written partly in response to a review of the first edition of *Critique of Pure Reason* (CPR)<sup>1</sup> which had presented transcendental idealism as rather similar to Berkeley's philosophy, Kant explained that his idealism is formal and not material.<sup>2</sup> This was carried over to a note in the B-edition of CPR, in which he contrasts his own idealism with the "common" material idealism that "doubts or denies the existence of external things" (B519).

The notion of formality employed by Kant to characterize his idealism is closely connected to that of *a priori* conditions of experience. Time and space are forms of intuition, the categories are forms of thought. The forms of intuition and the forms of thought together are conditions for objective experience. Formal or transcendental idealism is thus an idealism concerning the conditions for objective experience or cognition.

Henry Allison has stressed the importance of the distinction between transcendental idealism and transcendental realism for the understanding of Kant's philosophy.<sup>3</sup> Though Kant refers to transcendental realism only on a few occasions (A369–72; A491/B519; A543/B571), he does so in terms indicating that this position makes it impossible to maintain an empirical realism (a disastrous consequence according to Kant), and so leads to a material idealism (which is either skepticism about or denial of the existence of outer things) (B274; B519n). Transcendental realism and transcendental idealism are "mutually exclusive and exhaustive meta-

---

<sup>1</sup> References to CPR are according to the standard convention: "A" before the page numbers of the first edition (1781), "B" before those of the second edition (1787). The translation is that of Guyer and Wood (Kant 1997a). References to other works by Kant are given to the volume and page number of the *Akademie* edition (Ak.).

<sup>2</sup> Ak. 4:337; 375.

<sup>3</sup> Allison 1983, 14–30.

philosophical alternatives.”<sup>4</sup> They are exhaustive alternatives because there are just two possibilities regarding the distinction between things as appearances and considered in themselves: transcendental realism does not make this distinction, whereas transcendental idealism does.

Before spelling out the nature of this distinction, something needs to be said as to how so many different philosophies could possibly be classified as belonging in a common group. Transcendental realists, according to Kant, do not make the relevant distinction between things as appearances and things in themselves; when they make a distinction in these terms, they refer merely to the difference between illusory and veridical perception. Specifically, their mistake is that they (implicitly) hold that time and space pertain to things in themselves; Newton’s view of time and space as separately existing, and Leibniz’s, according to whom they are relations between things, are important examples of transcendental realism. Even if Kant refers only to the forms of intuition in this context, it is natural to extend the notion of transcendental realism to include also a view according to which the categories are predicates applicable to things in themselves.<sup>5</sup> Since appearances for Kant are objects in so far as they conform to our conditions for cognition, the transcendental realists, taking these conditions as belonging to objects apart from any cognition (an affirmation of the independent reality of that which for Kant is merely transcendental, that is, relative to the cognitive capacity), are in that sense conceiving of appearances as things in themselves (compare A369). The crucial point is that transcendental realism has no room for a transcendental level. According to such a view, there are no specifically transcendental or epistemic conditions<sup>6</sup> for our cognition of objects. Very roughly, empiricists take the representations of the senses to be things in themselves without any necessary reference to conditions for cognition, whereas rationalists, viewing the objects of the senses with some suspicion, regard

---

<sup>4</sup> Allison 1983, 14

<sup>5</sup> See Allison 1983, 28.

<sup>6</sup> The notion of an epistemic condition is from Allison 1983, 10–13. It is contrasted to logical, ontological and psychological conditions, and corresponds more or less to what Kant calls conditions of the possibility of experience. I will use the term as equivalent to “transcendental condition.”

## *The Thing in Itself*

knowledge of things as possible *a priori*, apart from any epistemic conditions. Positions very different in other respects can thus share this assumption concerning the status of the objects of cognition. Now it can also be seen why transcendental realism is supposed to engender skepticism (A491/B519). If things in themselves are taken to be available only indirectly, as hidden causes behind the representations of our senses, then it is hard to explain how we could ever get to know that they really exist. One option for those unsatisfied by this opportunity for skepticism is to abolish the thought of independent causes external to representations and allow only representations, now considered as actual things. These things would then be conceived not as epistemically conditioned, but as things in themselves. In this way even a radical empiricist can be taken to be a transcendental realist, and thus Kant is able to say that Hume took the objects of experience for things in themselves.<sup>7</sup>

As some use has already been made of the phrase “thing in itself,” its signification must now be discussed. It is customary to distinguish two main lines of interpretation: the two-world and the two-aspect approach. The two-world approach, sometimes called the textbook view, construes Kant’s distinction between things in themselves and appearances as a metaphysical theory according to which the real world (the thing in itself) is forever hidden beyond the reach of human knowledge, but gives rise to another world, the world of appearance, as a result of the affection of the things in themselves upon a subject. When the subject is affected by things in themselves, a process is initiated in which the forms of intuition and the categories are imposed on the material received, whereby appearances are constructed. This world of appearances consists of “mere representations” (a locution often used by Kant). It is not the real world, inasmuch as reality is taken to pertain to the things in themselves, but it is the only one we have access to, and what we call empirical knowledge relates to it.

Crucial for this type of interpretation is that thing in itself and appearance are two separate entities. Much of Kant’s thinking is seen as a difficult (not to say hopeless) attempt to deduce from the features of experience how the things in themselves are constituted. Because of the skeptical connotations inherent in the designation of the thing

---

<sup>7</sup> See *Critique of Practical Reason*, Ak. 5:53.

in itself as unknowable reality and experience as mere appearance, there is a tendency in two-world interpretations to reduce the world of experience to nothing more than a kind of illusion. In this picture, the world consisting of appearance “only appears to exist, is really nothing apart from perceptions.”<sup>8</sup> When pushed in this direction, the two-world interpretation only recognizes one world, the world of things in themselves, and apart from that a distorted quasi-world consisting of our perceptions, from which no information whatsoever can be gained concerning the real world in itself, despite Kant’s frequent attempts to obtain such impossible metaphysical knowledge. That the two worlds tend to collapse into one when the view is developed in a consistent way is confirmed by James Van Cleve’s summary of his own interpretation: “I ascribe to Kant a two-worlds view. Of course, since appearances are only virtual objects, there is a sense in which they do not really exist and so do not form a second world. But they do at least form a second domain of discourse, not simply a second way of talking about things in the first domain.”<sup>9</sup>

The main merit of this reading is that Kant makes quite a few statements that seem to imply a two-world view. To take just one passage, consider the following: “[the understanding] thinks of an object in itself, but only as a transcendental object, which is the cause of appearance (thus not itself appearance)” (A288/B344). Here, it seems, we have a thing in itself that causes appearance without itself being appearance, just as the two-world model requires.

There are both textual and philosophical inconveniences with this view, however. There are many passages where Kant speaks of the same things as appearances and considered in themselves, so that we seem to have one world regarded from different perspectives. For instance at Bxxvii he describes his “distinction between things as objects of experience and the very same things as things in themselves” as an important tenet of the *Critique of Pure Reason*. In the same place he also states that the *Critique* takes the object “in a twofold meaning,” namely, as appearance and as thing in itself. The fact that Kant often expresses the distinction in similar ways, as concerned with one and the same world, seems to go clearly against the two-world model. It is also difficult to understand how Kant could

---

<sup>8</sup> Strawson 1966, 238.

<sup>9</sup> Van Cleve 2003, 194–195. Compare Van Cleve 1999, 150.

attempt to describe causal processes going on behind the appearances, as the two-world view takes him to do, since causality is one of the epistemic conditions identified in Kant's theory, and as such is valid only inside the sphere of experience. He might of course *believe* that such processes take place, but it could not be part of his critical doctrine. Furthermore, it is not easy to see how the notion of unknowable things affecting us and thereby producing a world of appearance differs in any decisive way from the skeptical idealism that Kant attributes to transcendental realism, which holds that external things cause us to have representations, though we have no way to know whether our representations adequately represent these things.<sup>10</sup> The two-world approach with its skeptical implications also goes very much against Kant's confidence in the adequacy of natural science as cognition of the world, and his defence of empirical realism. It might be objected that the difference between this version of transcendental realism and Kant's transcendental idealism is only a shift of emphasis. Whereas the former takes the unknowability of things in themselves to imply that we have no genuine knowledge at all, Kant's point would then be that we should instead adopt the attitude that assertions about what is given as representations should be treated as real knowledge, perhaps in the form of a phenomenalist theory. This, I think, is how the two-world interpretation must deal with Kant's distinction between transcendental realism and transcendental idealism. But it is not very plausible to construe Kant's defence of empirical realism as only superficially different from skepticism, whether the difference is taken as verbal or psychological.

Problems such as these have prompted commentators to look for other ways of understanding the notion of the thing in itself.<sup>11</sup> Aiming at a philosophically more defensible interpretation of Kant, the two-aspect approach takes such passages seriously which speak of two

---

<sup>10</sup> A similar point is made in Bird 1962, 23.

<sup>11</sup> "Es wird wohl für immer eine der Merkwürdigkeiten der Kant-Forschung bleiben, dass jemals ein solches Konglomerat aus schlechter Metaphysik für das letzte Wort der Kantischen Philosophie gelten konnte und offenbar immer noch gilt", says Gerold Prauss (1977, 194) about one version of the two-world interpretation.

### *The Thing in Itself*

perspectives on the same thing.<sup>12</sup> According to Prauss, the phrase “thing in itself” is an abbreviation of “thing considered in itself” (*Ding an sich selbst betrachtet*), which corresponds to such Latin expressions as “*res in se spectata*” and “*res per se considerata*” used in the rationalist tradition.<sup>13</sup> The words “in itself” are not modifying “thing” (as if singling out a certain class of things); they are used to refer to a kind of consideration, a perspective on things as isolated from our cognitive conditions or more generally from any relations. On this view, the consideration in question is a process of abstraction, in which the *a priori* forms of sensibility (and in some contexts the *a priori* conditions of thought as well) are removed from the thought of the objects.

Support for an interpretation along these lines can be found in passages such as this: “If we abstract from *our* way of internally intuiting ourselves [...] and thus take objects as they may be in themselves, then time is nothing” (A34/B51). The objects “as they may be in themselves” accordingly are these same objects considered in abstraction from a certain epistemic condition, our way of intuiting. Admittedly, there is a certain difficulty in saying that we conceive of the same object in two ways, since the consideration of things in themselves by abstracting from the forms of intuition removes the condition for individuation of objects, which on Kant’s view is bound up with space and time. Given the suspension of a basis for the differentiation of objects, it is not possible to claim to know that the things in themselves constitute a plurality corresponding to the empirical objects, though it may sometimes be useful to describe the situation in that way. In his moral philosophy Kant thinks of the human being as retaining its individuality when considered in itself (as noumenon), and though this thought is problematical from a theoretical point of view, Kant finds it motivated from an internally practical perspective. But even when individuation is taken as inapplicable to the thing in itself, so that it is

---

<sup>12</sup> Influential works representative of this approach are Prauss 1977 and Allison 1983. But it has been proposed many times before: see for instance Heidegger 1991 [1929], 31–34, Bird 1962, Matthews 1969 and Phalén 1976, 87–89.

<sup>13</sup> Prauss 1977, 20.

considered just as the world in abstraction from epistemic conditions, there is no need to posit this world as an ontologically separate entity.

According to the two-aspect approach, the distinction between appearance and thing in itself is not ontological but epistemological or perspectival. To be sure, Kant sometimes uses the term “thing in itself” ontologically, to refer to hypothetical objects entirely outside of our cognitive reach (such as God), but the epistemological use is the primary one.<sup>14</sup> Some of the traditional problems of interpretation can be dissolved in adopting this view, for instance the question of how things in themselves can influence us “before” experience and thereby cause us to construct a world of appearances. There simply is no separate world of things in themselves that affect us in this way; there are things that appear to us, and there is a specifically philosophical consideration of these things in abstraction from the conditions of cognition. Even so, affection can be described in two ways: as an empirical process (perception), or, in abstraction from the epistemic conditions, as a noumenal fact. To represent the abstract, noumenal description of affection as a hidden process causally connected to and numerically different from what goes on in the empirical process of perception is a mistake that results from an ontological interpretation of a perspectival distinction.<sup>15</sup>

Likewise, the classic objection that Kant has no ground for claiming to have knowledge of the existence of things in themselves, considering that they are also said to be unknowable, loses much of its force on this interpretation. The two-aspect view does not introduce a new set of objects, the existence of which could be challenged. The Kantian approach is to accept that things exist, and to assume that they don’t cease to exist just because they are considered in themselves. When Kant dissociates his view from material idealism, he typically emphasizes the existence of *things*, not of things in themselves. The existence of the latter would need to be

---

<sup>14</sup> Meerbote 1974.

<sup>15</sup> In the context of a more sophisticated ontologically oriented account, Piché (forthcoming) helpfully construes affection in terms of two aspects of the same event. Compare Weldon 1958, 254: “Indeed ‘affection’ of the non-empirical self by the non-empirical object and the ‘affection’ of my empirical sense by the phenomenal object are actually the same event viewed from different standpoints.”



argued only if the *an sich* constituted a separate realm of being. Thus, in *Prolegomena* we find affirmations of the realism of Kant's transcendental idealism: "the existence of the thing that appears is not thereby nullified, as with real idealism."<sup>16</sup> This is not to deny that Kant sometimes expresses himself as if he were ascribing existence to separate entities called things in themselves: "the understanding, just by the fact that it accepts appearances, also admits to the existence [*Dasein*] of things in themselves."<sup>17</sup> Judged by strict two-aspect standards, this verges on the nonsensical. The phrase "things in themselves" is used as if it were the name of a kind of object. But it is not difficult to take the sentence to mean that the understanding, by accepting the existence of things as appearances, thereby is committed also to the existence of these very things considered in themselves, apart from our cognitive conditions (though this implies no claim that we are able to give any positive description of them in this respect). This might now seem to be a trivial point, perhaps too trivial to assume it to be the one Kant wants to make. But consideration of the argumentative background gives some explanation as to his motivation. In the *Prolegomena* Kant tries to counterweigh the subjective idealism ascribed to him in Garve's and Feder's review of the first edition of the *Critique of Pure Reason*. He takes every opportunity to stress the realistic strand in his philosophy, already firmly stated in the doctrine of empirical realism. One way for him to do this is by denying any commitment to the non-existence of things in themselves.

No definite unanimity can be expected as to how to judge the relative importance of two-world sounding passages as against the ones that suggest a two-aspect frame.<sup>18</sup> Kant's basic commitment to realism (which is an expression of his empirical realism) and the two-aspect view are compatible on the above reading of the passage from *Prolegomena*, which at least suggests that what may look like an appeal to a two-world ontology need not always be that.<sup>19</sup>

---

<sup>16</sup> Ak. 4:289; compare 4:293, and also B519n.

<sup>17</sup> Ak. 4:315.

<sup>18</sup> See Van Cleve (1999, 144–146) for an assessment differing from the present one.

<sup>19</sup> Incidentally, just before the sentence discussed, Kant employs the locution of one thing viewed in two different respects: "we are not acquainted with

### *The Thing in Itself*

Hoke Robinson has objected to Allison's two-aspect view on the ground that two ways of considering something "must ultimately deal with the *aspects* to which the considerations are directed."<sup>20</sup> A thing can be considered in different ways, but not in just any way; and the different aspects have to be consistently referrable to the same object. According to Robinson, Allison is not able to account for the sameness of the object that has the two aspects. Allison has replied that Robinson's critique is based on the assumption that the different aspects have to correspond to different pre-given aspects in the underlying thing.<sup>21</sup> But this, Allison argues, is to miss the point: the two-aspect view is committed to the possibility of two distinct epistemic relationships in which things can be considered, in conjunction with the thesis that spatio-temporal properties pertain to things only in relation to one of these considerations. This is convincing as a rebuttal of an overly ontological interpretation of Kant, but Allison's solution to the question of sameness is less satisfying. The problem is "that these two conflicting ways of considering do not seem to leave room for the conception of some single underlying thing that is being considered."<sup>22</sup> As the two aspects are two points of view, neither of them can serve as the neutral basis for the characterization of that which is considered in the two ways. To solve the problem, Allison assigns this function to what Kant calls "the transcendental object=x."<sup>23</sup> Though it is true that "the transcendental object" in CPR refers to the indeterminate concept of a something in general, and as such is somewhat more neutral than a positive conception of a thing in itself as thought by the pure categories, it is not clear that there is a problem here that needs to be solved. The sameness of the thing under different considerations is guaranteed by the internal consistency of the reflection in which the different considerations take place. As long as the considerations do not ascribe contradictory predicates to the same thing – and Allison

---

this thing as it may be constituted in itself, but only with its appearance" (Ak. 4:314–315).

<sup>20</sup> Robinson 1994, 421f.

<sup>21</sup> See Allison 1996, 12–17.

<sup>22</sup> Allison 1996, 16.

<sup>23</sup> Allison 1996, 16.

### *The Thing in Itself*

explicitly claims that they do not<sup>24</sup> – there is no need for a neutral description of the single underlying thing. The considerations are from the very start defined as being of the same object,<sup>25</sup> and if we must say *what* this object is, we might say it is the object of appearance, since that is the one we ordinarily cognize. If we wish to describe this on a transcendental level, in abstraction from cognitive conditions, we could say that the object in question is a transcendental object, but we are not forced to find a way of referring to the object independently of the reflection or consideration from which we started. In the context of the philosophy of action, Anscombe has argued that it is a mistake to think that we have to use a non-descriptive designation in order to be able to assign different descriptions to the same act.<sup>26</sup> Analogously, the two-aspect theorist should reject the demand for a neutral designation of the thing reflected upon.

Let us now take a closer look at a passage that was quoted above, where Kant talks of a thing in itself as causing appearances, on the face of it suggesting an ontological distinction. The understanding “thinks of an object in itself, but only as a transcendental object, which is the cause of appearance (thus not itself appearance)” (A288/B344). To understand this, we need to be more precise concerning what it is that is removed when a thing is considered in itself, and for this purpose I shall first briefly summarize Kant’s theory of sensibility and understanding. The two-aspect view was

---

<sup>24</sup> Allison 1996, 16.

<sup>25</sup> To some extent this can be compared to Kripke’s view that when we speak about a thing in a counterfactual situation, the identity of the thing spoken about is simply “*given* in the very description” or stipulated (Kripke 2001, 229). The analogy should not be pressed too far though, since Kripke’s discussion concerns very different issues.

<sup>26</sup> Compare Anscombe 1979, 220: “I have on occasion stared dumbly when asked: ‘If one action can have many descriptions, what is *the* action, which has all these descriptions?’ The question seemed to be supposed to mean something, but I could not get hold of it. It ought to have struck me at once that here we were in ‘bare particular’ country [...]. The proper answer [...] is to give one of the descriptions. Any one, it does not matter which; or perhaps it would be best to offer a choice, saying ‘Take whichever you prefer.’”

introduced above as saying that a thing in itself is a thing considered in abstraction from transcendental conditions. These conditions, however, are of two separate kinds: the sensible forms of intuition (space and time) and the intellectual forms of thought (categories). Objective experience is achieved in judgments. The understanding judges by applying concepts (general representations) to particulars. In empirical cognition, these particulars are given in intuition, and the judgment is therefore a unification of the contents of intuition by means of concepts. In this interplay, sensibility is the passive recipient of the material that the understanding actively coordinates (A50/B74). Our understanding is discursive, which means that it cannot provide the objects or intuitions (singular representation) that are needed to give content to its general concepts (see A68–69/B93–94). So, as a consequence of our cognitive constitution (not by logical necessity), we need the material of the senses as a condition for cognition. But a further condition is that this material is received in accordance with the forms of intuition. No explanation can be given for the fact that we have space and time as our forms of intuition, rather than some other medium of receptivity; this might be called a “transcendental fact” (compare B146). By this expression is meant that it is neither an empirical fact (since it is a condition for the possibility of the empirical as such) nor a logical necessity (since it is not incoherent to think that other beings could have other forms of intuition). The peculiar character of this transcendental level, distinct from logic as well as from the empirical level, is expressed in another context in the doctrine of synthetic judgments *a priori*, which are synthetic like empirical judgments, and *a priori* like logical principles, but distinct from both. We can frame the thought of beings with other forms of intuition, though we cannot imagine what such forms of intuition would be like. As for the forms of the understanding, on the other hand, Kant’s position appears to be that no other logical forms (forms of judgment) can be conceived than the ones that we have, not even as a mere abstract possibility, if the understanding is to be discursive.<sup>27</sup> On the other hand, we can con-

---

<sup>27</sup> At the end of § 21 in the B-Deduction (B145–146) there is some ambiguity in this regard: that we have precisely the “functions for judgment” that we have is said to be impossible to give a further ground for, as it likewise is impossible to explain why space and time are our only forms of intuition.

ceive of an intuitive (non-discursive) understanding, radically different from ours, whose thinking would also be an intuiting.

Consequently, if there were beings endowed with other forms of sensible intuition than ours, they would have to use the same logical forms as we do to attain cognition (B148). The reason for this incongruity between forms of intuition and logical forms is perhaps that the latter amount to what Kant takes as a complete set of combinatory possibilities, namely in regard to judgments of subject-predicate form.

We are now in a position to explain a more specific sense of “thing in itself” often used by Kant. The thing in itself need not be the thing considered apart from *all* epistemic conditions; it can also be the thing considered apart from the sensible conditions only. Things in themselves are then thought by the understanding, but as devoid of sensible content. When Kant criticizes Leibniz for intellectualizing appearances (A271/B327), he points out that things for Leibniz are things in themselves in this sense. This criticism is to some extent applicable also to Kant’s own *Inaugural Dissertation* of 1770, where he defines intellectual cognition as referring to things as they are, whereas sensible cognition is about things as they appear.<sup>28</sup> The rationalistic ideal of a purely intellectual knowledge is one important example of what Kant later called transcendental realism, and the objects of this knowledge were supposed to be things in themselves, things in abstraction from all sensible conditions.<sup>29</sup> In the B-edition of CPR, Kant introduces the term “noumenon in negative sense” for the thought of an object in abstraction from our mode of intuition, which he describes as an entirely indeterminate something outside of our sensibility, in distinction to a positively conceived noumenal object as

---

This might suggest the possibility of alternative kinds of discursive understanding, but should presumably only be taken to counter the demand for an ultimate explanation of the conditions for cognition.

<sup>28</sup> Ak. 2:392. In the *Dissertation*, however, an independent status for cognition by the senses is accorded.

<sup>29</sup> Patt (1987, 20) even sees this as the original sense of “*Ding an sich*” in the A-edition: “‘Ding an sich’ ist der Ausdruck, mit welchem die kritische Philosophie die Leibnizsche Ding-Auffassung kennzeichnet.” Allison, modifying his earlier view, now sees the consideration of things through the mere understanding, in abstraction from sensible conditions, as the more precise notion of the thing in itself. See Allison 1996, 7–8.

object for an intellectual intuition (B307). In some contexts, the noumenon in negative respect replaces the A-edition's notion of the transcendental object. One can surmise that the reason for this terminological adjustment was to simplify matters, as well as to avoid confusing two rather different uses of "transcendental object," namely the one just discussed, and another one in the *A-Deduction* according to which "the transcendental object" designates the concept of the unity of representations, correlated to the transcendental unity of apperception, so that the term in this context seems to function somewhat like the concept of substance.<sup>30</sup>

To return now to the passage quoted above, which says that the understanding "thinks of an object in itself, but only as a transcendental object, which is the cause of appearance (thus not itself appearance)" (A288/B344), it can be noted that it is the understanding that thinks this "object in itself," and as it merely thinks it, but has no corresponding intuition of it, it is only a "transcendental object", which suggests that this term is here equivalent to the "noumenon in negative sense". This is conceived as the cause of appearance and therefore not itself appearance. I contend that no hypothesis about an ontologically separate thing in itself causing appearances is proposed here by Kant.<sup>31</sup> He is rather describing how the understanding, in abstracting from sensible conditions, uses its pure logical forms to conceive of a something in general, and how it conceives of this something as the ground of appearances merely because something remains after the abstraction, namely the pure thought of an indeterminate object. Conceiving this as the ground of the object amounts to thinking that the object is not annihilated when the sensible conditions are cancelled. The passage occurs in the context of a discussion of the limitation of sensible cognition, and this limitation is marked precisely by the thought that something remains even if space and time are thought away. In this reading, "ground" is used instead of "cause" (in accordance with Kant's explanation in the *Critique of the Power of Judgment*: "the word cause, when used of the supersensible, signifies only the

---

<sup>30</sup> See the discussion in Klemme 1996, 245–277, as well as Allison 1983, 244, Buchdahl 1989, 224, and Robinson 2001.

<sup>31</sup> Whereas Van Cleve (1999, 162) takes the passage as support for the view that things in themselves affect us.

ground”<sup>32</sup>) in order to stress the purely logical character of this use of the understanding. Ground, as the logical form corresponding to the category cause, is the more appropriate term for a thought which abstracts from the sensible condition required for the employment of a category (cause) in empirical cognition. Though Kant sometimes uses the word “cause” in a somewhat careless way, reading it as signifying “ground” suggests that he does not bring in a causal relation between two things, one noumenal and one phenomenal, but rather expresses the point that no ontological consequences can be drawn from the methodological abstraction involved in transcendental philosophy.

Whatever one thinks of the proposed interpretation, it should at least show that a reading in terms of two worlds is not the only possibility. But if the two-aspect interpretation is a plausible reading of the text, a different question poses itself. What possible interest could such a view have? It appears to be threatened by triviality. It is not very controversial to say that things as considered in abstraction from the conditions for knowledge cannot be known. What prevents the Kantian view from collapsing into such emptiness is its theory of transcendental conditions.<sup>33</sup> The thesis of the unknowability of things in themselves gets its significance from the concomitant theses that space, time and the categories are *a priori* forms of cognition, and that these conditions determine the formal aspects of experience. The distinction between appearance and thing in itself serves to articulate transcendental idealism as a theory of the cognitive faculties and their role in contributing *a priori* forms essentially related to the kind of subject possessing these faculties.

Rae Langton challenges Allison’s interpretation of things in themselves on the ground that it collapses into mere emptiness.<sup>34</sup> According to her, it amounts to the following anodyne thesis:

---

<sup>32</sup> Ak. 5:195.

<sup>33</sup> Compare Strawson 1994, 170–171. Here a significant step is taken beyond the position of *The Bounds of Sense* (Strawson 1966). Whereas he then saw the issue of things in themselves and transcendental idealism in terms of a two-world model, Strawson now accepts the two-aspect interpretation as a possible alternative, though he finds the two-world model more in conformity with Kant’s text.

<sup>34</sup> Langton 1998, 9.

### *The Thing in Itself*

Things considered in abstraction from their relation to our sensibility are things considered in abstraction from their relation to our sensibility.

This tautology is obtained from Kant's claim that we can have no knowledge of things in themselves (presumably by substituting for "knowledge" what according to Kant is a necessary condition for knowledge, namely "relation to our sensibility," and for "in themselves" the two-aspect reading of that locution, "considered in abstraction from their relation to our sensibility").<sup>35</sup> Now, as Langton remarks, Kant thinks he has made a philosophical discovery.<sup>36</sup> How could that be the case if the unknowability of things in themselves is just a tautology?

The situation can profitably be compared to Frege's well-known astronomical example in his discussion of sense and reference. Ancient astronomers believed that the Morning Star and the Evening Star are different heavenly bodies. Let this belief correspond to the pre-Kantian philosopher's belief that we have knowledge of things by means of the intellect, unaided by sensibility. Then, in the astronomical story, someone discovers that the Morning Star and the Evening Star are one and the same planet. In our story, Kant claims that he has discovered that we have no knowledge of things by means of the pure intellect, apart from the conditions of sensibility. Both discoveries are, in their respective contexts of discourse, informative and perhaps controversial. However, in both cases, from the new point of view attained, the discovery can be construed as a mere tautology. In Frege's story, the astronomer's assertion that the Morning Star is the Evening Star can, by substitution, be transformed to the claim that the Morning Star is the Morning Star. In Kant's case, his claim that we can have no knowledge of things in themselves can be transformed to Langton's tautology. Frege discharges the paradox with the help of his distinction between sense and reference. Likewise, it is possible to view our problem in a somewhat Fregean

---

<sup>35</sup> The derivation of the tautology is motivated thus: "Since knowledge arises only with the concrete application of the forms and categories, Kant's thesis [...] about our ignorance of things in themselves becomes nothing more than this [tautology]" (Langton 1998, 9).

<sup>36</sup> Langton 1998, 10.



spirit (irrespective of what the ultimate value of the sense-reference distinction may be in semantics). In an intellectual community dominated by the belief (shared by Kant in his *Dissertation* of 1770) that we have knowledge by means of the intellect of things as they are independently of their relation to our sensibility, the denial of this belief is a substantial thesis, even if this denial, supplemented with the new theory that empirical knowledge is conditioned by the forms of intuition, may yield a tautology. I say “may”, since it is not clear to me that there is a tautology. Langton obtains one by substituting for “knowledge” the Kantian transcendental conditions (forms of sensibility and categories) in “We can have no knowledge of things in themselves”. Supposing that Kant’s argumentation for the requirement of these conditions is correct, perhaps they can be said to be analytically linked to the concept of human knowledge. But the question Kant faced in developing his critical philosophy was precisely whether there are such conditions for human knowledge, or whether it is based on intellectual intuition, a condition-free immediate cognition. To answer that question was not a matter of analyzing the existing concept of “knowledge,” but rather of forging a new conception of human knowledge.

Another way to make this point is to ask who are included in the “we” of the thesis “We can have no knowledge of things in themselves.” If it is all knowers in general, the thesis is no tautology according to Kant (for perhaps God has such knowledge). If only beings are included which have a twofold cognitive structure, in which the application of concepts (the understanding) is dependent on reception of data (sensibility), then the thesis may be tautological, but the question whether human cognition is of this dual kind is a substantial one, which Kant, contrary to empiricists and rationalists (assuming the correctness of his interpretation of these positions), answered affirmatively.

Moreover, it seems possible to construe just about any interpretation of the unknowability of things in themselves as tautological in the way indicated by Langton. For instance, in her own interpretation of Kant, according to which “in itself” designates the intrinsic (non-relational) properties of things, the unknowability thesis becomes “We have no knowledge of the intrinsic properties of

things.”<sup>37</sup> Kant’s view is then that our knowledge, being the result of things affecting us, is necessarily relational, since “the causal powers of things do not supervene on their intrinsic properties,” so that “it is not through those intrinsic properties that we are affected.”<sup>38</sup> But then we can replace “knowledge of” with “relation to” (in parallel to how Langton, in her construal of Allison’s interpretation, replaced “knowledge” with the relation to sensibility that for Kant is the transcendental condition for knowledge), thereby obtaining “We have no relation to the non-relational properties of things” (since intrinsicity is non-relationality). This certainly looks like a tautology.

The proper way to deal with such a charge of emptiness is to stress that the claim that our knowledge is necessarily relational is a substantial one, and that consequently Langton’s reading of Kant is in no way trivial. Similarly, the claim that our knowledge is based on transcendental conditions is not trivial, and it is therefore pointless to construe the two-aspect reading as a tautology.

## 2. METAPHYSICAL AND METHODOLOGICAL VIEWS

Having said this in defence of the two-aspect view, it must be admitted that the issue goes somewhat deeper than it has been depicted so far. First, there is the linguistic point that the word “world” does not necessarily mean an ontologically separate realm but sometimes just indicates a general aspect of things, so that things considered apart from epistemic conditions could well be referred to as a world. To find that an interpretation explains Kant’s distinction in terms of “worlds” is therefore not sufficient to determine that it is an instance of the two-world view. Secondly, and more importantly, recent accounts sympathetic to the two-world approach tend to be more sophisticated than earlier versions. Rather than the question of one or two worlds, the recent debate concerns whether a “metaphysical” or a “methodological” interpretation of the thing in itself is

---

<sup>37</sup> Langton 1998, 13.

<sup>38</sup> Langton 1998, 139. I will have more to say about Langton’s interesting proposal below.

the most adequate one.<sup>39</sup> Whereas a methodological view considers Kant's contrasting conceptions of the noumenal and the phenomenal in terms of different reflective perspectives or epistemic points of view prompted by the methodology of transcendental philosophy,<sup>40</sup> the metaphysical interpretation distinguishes two sets of properties belonging to objects, noumenal and phenomenal properties.<sup>41</sup> The noumenal properties are the intrinsic properties of the object, whereas the phenomenal ones are properties that the object possesses in its relation to the knowing subject with its forms of intuition. In construing the distinction of thing in itself and appearance as based on different sets of properties of the same object, the metaphysical view takes the form of a two-aspect theory. There is a deeper resemblance to the two-world view in that the existence of unknowable properties belonging to the thing in itself is taken to be established by means of general metaphysical reasoning, largely independently of the specifics of Kant's transcendental methodology. It is true that at least formally, the metaphysical view amounts to a one-world view. It postulates only one world, one set of objects that have properties belonging to these objects as things in themselves as well as properties that are phenomenal. Nonetheless, this approach can justly be associated with the two-world theory, since both views take Kant's distinction between appearance and thing in itself to be primarily ontological rather than epistemological. The metaphysical view takes the *an sich* to consist in special properties, whereas the two-world view takes it to be a realm of separate things. What both views have in common is that Kant's distinction is considered to involve the assumption of a realm of entities of different ontological status than objects on the empirical level have. A metaphysical view assumes what might be called a *transcendent* point of view, even when it accepts Kantian strictures on our ability to attain cognition

---

<sup>39</sup> See Westphal 1997, 232–233 and Grier 2001, 86 and 89 for characterizations of the two positions.

<sup>40</sup> Versions of the methodological approach in the more recent literature are found in Buchdahl 1992, Palmquist 1993, Rescher 2000 and Grier 2001. Kroon 2001 offers a hybrid view.

<sup>41</sup> The metaphysical interpretation is defended in different versions by Westphal 1997, Langton 1998, Van Cleve 1999 and Langsam 2000. Adams 1997 can also be counted to this group, though in some respects his view resembles that of the methodological interpretation.

transcending experience, and it considers it legitimate to some extent to pursue metaphysical investigations regarding the nature of the thing in itself. In so far as it is granted that “thing in itself” and “appearance” refer to two aspects of the same thing, the *real* being of this thing is taken to be the thing in itself, while appearance is taken to be some particular, derivative property of the thing in itself. The metaphysical view takes a top-down approach in that it assumes it possible to determine the ontological status of things in themselves and appearance.

It can already be seen from this sketchy presentation that Robinson’s objection to Allison, referred to above, is framed from a metaphysical view of things in themselves, and that Robinson presupposes that Allison’s two aspect-view must ultimately be metaphysical in this sense. What Robinson demands is a metaphysical characterization of an entity underlying cognition. Though Allison’s response is somewhat vacillating, I take it that he attempts to defend a methodological view of things in themselves that, as it were, proceeds from inside of experience. In such an immanent view,<sup>42</sup> metaphysical theorizing is taken as valid only when conducted in a *transcendental* frame, that is, in so far as it springs from an explication of *a priori* structures of experience. In so far as *transcendent* metaphysical speculation is allowed also in this approach, it must be in a bottom-up direction, so to speak as an afterthought, rather than as the primary move. In addition, metaphysical themes have a special role in moral philosophy, where they are considered by Kant to serve more than merely heuristic functions, but yet only from a practical point of view.

Proponents of the metaphysical view often allege that the methodological view is unable to accommodate metaphysical possibilities of obvious importance for Kant’s account of moral philosophy, such as the freedom of will. Some of the resistance to the methodological reading may be caused by a feeling that it is too narrowly concerned with empirical cognition. But it is not clear that the methodological interpretation must lack all resources for openmindedness in metaphysical questions. A methodological inter-

---

<sup>42</sup> See A295–296/B352–353 on immanent principles (though my use of “transcendent” and “transcendental” here differs from Kant’s use in that passage).

pretation of the thing in itself at least admits metaphysical *possibilities* that may legitimately be reflected on from a practical perspective, if not in theoretical cognition.

Moreover, on this interpretation the thing in itself introduced in the *Transcendental Aesthetic* can unproblematically be claimed to exist, because it is the same thing as the empirical object conceived of in abstraction from conditions of sensibility, as was discussed above. Hence the thing in itself, for all its unknowability, “inherits” its existence from the empirical realism accepted for ordinary empirical things, from which the process of abstraction started. This is sometimes overlooked when the absence of an ontological stance in the methodological approach is deplored.

### 3. THINGS IN THEMSELVES IN PERSPECTIVE

The two-aspect view introduces different reflective perspectives on the same object. It also brings in perspectival considerations about Kant’s potentially misleading use of the term “thing in itself” in empirical contexts. Sometimes, in describing ordinary perceptual situations, Kant calls an object a thing in itself, and contrasts it with the diverse subjective perceptions that can be had of this same object. The perceptual seemings are called appearances, the same term that is used in the transcendental distinction.<sup>43</sup> Though Kant makes it clear that this empirical use of the distinction contrasts with its transcendental use, the fact that he uses the same terms on different levels of reflection might lead a reader astray. Moreover, it is important to keep in mind that the relation between the terms differs on the two levels, at least on one construal of perception. In the empirical situation, my perceptual “appearance” is of the “thing in itself”, but it is distinct from it. An empirical thing in itself, for instance a rose (A29–30/B45), is not a perception considered in abstraction. Moving to a transcendental context, the thing, which was a thing in itself in the empirical respect (such as the rose), is an appearance in transcendental respect, and is contrasted with the thing in itself (in

---

<sup>43</sup> See A29–30/B45; A45–46/B62–63, and discussions in Meerbote 1974 and Prauss 1977, 47–51.

### *The Thing in Itself*

transcendental respect). But now the thing in itself is this thing (the appearance) considered in itself, in abstraction from epistemic conditions. Since the relation we find on the empirical level between two different entities, the perceptual appearance and the actual thing, is not reproduced on the transcendental level, the dissimilarity between the two cases is as great as the analogy. On another view of perception, we find a closer similarity between the empirical and the transcendental cases. If appearance is taken to be the perception of a thing from a certain perspective (as defined by the relevant physiological and sensorial factors), then my perception of the rose, though an appearance, is the rose itself perspectivally given. In parallel to this, appearance in transcendental respect is the thing as given from the perspective defined by the conditions of space, time, and the categories, whereas the thing in itself is the thing conceived of in abstraction from these conditions.

Kant distinguishes different levels of reflection, such as the empirical and the transcendental. Sometimes the same terms are used at different levels, so that we need to know at what place in the system we are in order to determine what they point to. And in certain contexts, even the terms indicating the levels can be rather flexibly used. For instance, in the introduction to the *Transcendental Dialectic*, it is said that the “transcendental use” of the principles of the understanding is their illegitimate use outside of experience (A296/B352–353), which is somewhat surprising; one should rather have expected this illegitimate use to be called “transcendent.”<sup>44</sup> Furthermore, Kant’s philosophical system consists of autonomous domains governed by different interests of reason (the practical and the theoretical) and we can expect that the interpretation of terms will vary relative to this. It might be said that a term such as “thing in itself” has a perspective-dependent meaning, but perhaps it is better to view it as a question of it being used to point out different aspects of things or ways of considering things relative to different reflective contexts, much like how the words “to the left” can point to rather different things from different vantage points, without having a different meaning each time.

---

<sup>44</sup> But Kant here reserves “transcendent” for principles that recognize no boundaries of experience.

### *The Thing in Itself*

Any interpretation of Kant is to some extent perspectival. But some are so more than others, and perhaps most of all those of Gerd Buchdahl (1989, 1992) and Stephen Palmquist (1993). I shall present some of their views on the interpretation of “thing in itself” and related topics (Buchdahl’s more in detail, Palmquist’s but briefly), putting them in the context of a methodological two-aspect approach.

Buchdahl presents what he calls a “flow chart of Kant’s transcendental dynamics”, a diagram in which the concept of an object is pictured as passing through various stations as a result of certain operations of “reduction” and “realization.”<sup>45</sup> In this highly reconstructive interpretation of Kant, shown in *Figure 1* below, the starting point is the object (or the world) in its “phenomenological” aspect, that is, as it is experienced before any philosophical investigation is conducted.<sup>46</sup> Buchdahl calls this  $T_w$  (object in the sensory world), and compares it to the Husserlian notion of the lifeworld; in a rather loose way, Buchdahl’s procedures are inspired by Husserl’s phenomenology. The first step in providing an “ontological” account of what is thus phenomenologically given is to *reduce* it in a process of abstraction (somewhat resembling Husserl’s phenomenological reduction), reaching  $T_g$ , corresponding to Kant’s *Ding überhaupt* (thing in general). By further abstraction we move to the stage  $T_o$  (what Kant calls the “transcendental object”). The precise relation between  $T_g$  and  $T_o$  is not entirely clear, but the difference is in terms of how far the abstraction from phenomenological qualities of experience proceeds. At stage  $T_o$  the object is considered as emptied of meaning or devoid of “ontological value.” Here reduction comes to an end, and a process of realization can take place. While reduction can be seen in terms of abstraction, realization is a constructive step, giving content or determining (a term often used by Kant) the reduced concept of the object. Such realization can be done in many ways, all not equally valid. According to this view, when Kant argues that space and time have the status of *a priori* forms of intuition, and that the categories are necessary conditions for experience, he tries to establish one kind of realization as legitimate. This is the realization of  $T_a$ , the object as appearance. Other realizations discussed by Kant

---

<sup>45</sup> Buchdahl 1989, 218.

<sup>46</sup> The following exposition is based on Buchdahl 1989. See also Leppäkoski 1993, ch. 5.

aim to reach noumenal stages. These realizations are characterized by their problematic status. Buchdahl construes Kant's distinction between the negative and the positive sense of the noumenon as two successive realizational steps. The noumenon in the negative sense ( $T_n^-$ ) "denotes a stage which represents  $T_o$ , as it were, already 'on the way to' [...] a potential realization, whether successful or not."<sup>47</sup> To reach something positive (a noumenon in the positive sense), there are two options, a theoretical or a practical realization. To make a theoretical realization of the noumenon ( $T_n^+$ ), we would need a capacity for intellectual intuition. Since we don't have that capacity,  $T_n^+$  at most has analogical or fictional sense for us. In the practical domain,  $T_n^+$  is realized, for there we consider ourselves in a noumenal perspective,<sup>48</sup> but this realization has no validity in the theoretical or speculative domain.

These are the stages in Kant's reduction-realization approach to the concept of an object in Buchdahl's reconstruction. But in his diagram there is also an alternative line of reduction-realization, going from  $T_g$  via reduction to a purely logical notion of an object ( $T_l$ ) to  $T_m$ , a metaphysical conception of the object achieved by a realization presupposing a capacity to cognize things in themselves. This line represents transcendental realism, a position which on Kant's view is common for pre-critical approaches to philosophy, in both the rationalistic and the empiricist forms. Transcendental realism regards cognition as realizable without transcendental conditions (such as space, time and the categories), which for Kant amounts to a conflation of appearance and thing in itself. The difference between  $T_g$  and  $T_l$  is rather unclear, but Buchdahl's inclusion of the latter in the diagram apparently aims to capture Kant's complaint that metaphysicians try to obtain knowledge of things merely by means of logical analysis of concepts. Buchdahl indicates that  $T_m$  in many ways resembles  $T_n^+$  (theoretical), and can be taken as a pre-critical version of it.

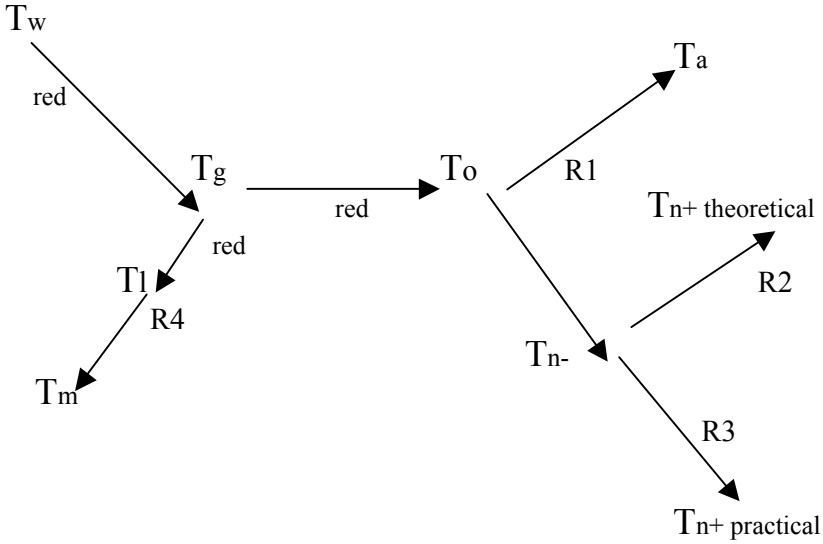
---

<sup>47</sup> Buchdahl 1989, 224.

<sup>48</sup> See for instance *Critique of Practical Reason*, Ak. 5:6.



*The Thing in Itself*



- R1 – Realization via sensibility and understanding (the human case).
- R2 – Realization via intellectual intuition (counterfactual).
- R3 – Practical realization.
- R4 – Metaphysical realization of the object, as assumed in precritical philosophy.

*Figure 1* Concepts of objects resulting from operations of reduction and realization. (Simplified version of the diagram in Buchdahl 1989, 218.)

Two terminologically interesting points can readily be spotted in Buchdahl’s model. First, there is no unique place in the scheme for the *thing in itself*. This locution is seen to be used by Kant for several different functions, corresponding to different stages in the scheme. In polemical use, when pre-critical views are criticized, it refers to  $T_m$ . Sometimes it denotes  $T_o$ , at other times the noumenon in one of its three modes:  $T_{n-}$ ,  $T_{n+}$  (theoretical), or  $T_{n+}$  (practical). Many of the problems in interpreting Kant’s view on “the” thing in itself are generated by insufficiently differentiating these various stages.<sup>49</sup> Second, though the term “appearance” is connected to  $T_a$ , it seems

---

<sup>49</sup> Buchdahl 1989, 226.

that the phenomenological starting-point  $T_w$  might also be called appearance, since in both cases we have to do with the empirical realm. And indeed, according to Buchdahl, Kant often uses the same term for these different concepts.

This prompts the question what “the point of all this R-R [reduction-realization] rigmarole” is.<sup>50</sup> Why would Kant perform a reduction of given experience only to realize it again? Buchdahl’s answer is perhaps not wholly clear, but its gist is that the constructive account in terms of realization is a way to explicate how space, time, and the categories can be *a priori*, by showing how these conditions function to make experience and cognition possible. The same situation is not uncommon in phenomenological (as well as other forms of) epistemology: cognition or experience is reduced or decomposed, and subsequently rebuilt again according to some methodological principle. Even though what is finally reached is the very thing from which one began, it is assumed that the process has provided a philosophical understanding not otherwise attainable.

Several objections to Buchdahl’s scheme are offered by Kenneth Westphal.<sup>51</sup> The main thrust of the objections is that Buchdahl’s tendency to construe Kant’s philosophy as a matter of choice is at odds with Kant’s general philosophical attitude. The following variations on this theme can be mentioned. (1) To make a choice, such as adopting a reduction-realization approach, we must be self-conscious; but, according to Kant’s argument in the Refutation of Idealism (B274–279), self-consciousness requires spatio-temporal identification of objects and events. Therefore, adopting a particular realization scheme in regard to objects cannot be a matter of choice.<sup>52</sup> (2) The R-R process amounts to a framework the adoption of which cannot be argued for independently. The view that justification can be given only within a framework is one that Buchdahl shares with Carnap. It is a problematic view, because it doesn’t allow us to ask for reasons to elect a particular framework rather than another one, though such a question ought to be legitimate.<sup>53</sup> (3) Kant’s strong

---

<sup>50</sup> Buchdahl 1989, 223.

<sup>51</sup> See Westphal 1998, and also Westphal 1997. For a defence of Buchdahl, see Jardine 2003.

<sup>52</sup> Westphal 1998, 342.

<sup>53</sup> Westphal 1998, 344.

modal claims for his doctrines are weakened in Buchdahl's account. Where Kant thinks his result is the unique solution of a problem, in Buchdahl's scheme it is merely a possibility.<sup>54</sup>

As concerns (1), apart from the general question of the appropriateness of construing Kant's doctrines on the conditions for experience as somehow being a matter of choice, the objection also focuses on the problematic assumption that a pre-given experience is available as a starting-point for reflection. This is problematic since Kant wants to establish conditions for the very possibility of experience. It would thus seem that either the R-R model is irrelevant to Kant's attempts to offer a philosophical explication of the structure of experience, or it must be supposed to represent conditions valid independently of (or processes having taken place in advance of) a particular person's choice to perform the abstractive and realizational operations described in the model.

The second alternative is preferable. It is incompatible with Buchdahl's aims only if his description of the procedure as entirely creative and constructive is taken at face value. Buchdahl tends to describe it in such voluntaristic terms, as a creation of a framework for the articulation of a certain philosophical position.<sup>55</sup> It seems to me that it is more in line with Kant's general attitude, and consistent with Buchdahl's interpretive scheme (if not with what might be called his ideological interpretation of this scheme) to consider the process of realization as guided by proofs and arguments, rather than as a free creation. It should thus be admissible to give a somewhat more traditionalistic interpretation of Buchdahl's scheme (whatever his own intentions might have been). The process of reflection depicted in the R-R model is then conceived of as a process of discovery, in which necessary conditions for experience corresponding to the realizational steps are isolated. Considered in that way, Buchdahl's perspectival model of Kant is to a large extent in agreement with Allison's two-aspect view.<sup>56</sup> It provides a pictorial view of the many facets of the concept of the thing in Kant.

---

<sup>54</sup> Westphal 1998, 346.

<sup>55</sup> See Buchdahl 1989, 224.

<sup>56</sup> They disagree on the role of transcendental object, which in Buchdahl's view must be sharply separated from the noumenon in negative sense.

### *The Thing in Itself*

Westphal's objections (2) and (3) can be treated in the same manner. Buchdahl tends to over-emphasize the constructive character of transcendental reflection, but it is more consistent with Kant's view to consider the goal of the reflective process to be the discovery of necessary conditions present prior to reflection. Bearing that in mind, Buchdahl's model remains useful as a graphic representation of the two-aspect interpretation.

A different and more detailed account of Kantian perspectives has been developed by Stephen Palmquist (1993). Palmquist outlines a framework of basic perspectives in Kant's system, supplemented with a set of reflective perspectives operative in each of the basic perspectives. He calls the first (basic) perspectives "standpoints." They are the theoretical, the practical, and the judicial standpoint. To each of these basic standpoints belongs a *Critique* (*Critique of Pure Reason*, *Critique of Practical Reason*, and *Critique of the Power of Judgment*). In each standpoint investigations can be conducted from different reflective perspectives, namely the empirical, the transcendental, the hypothetical, and the logical. Aiming at a general survey of Kant's entire architectonic, Palmquist describes how these perspectives are expressed in the basic standpoints. What is important with regard to the case of thing-terms discussed here is not the complex detail of his account, but just the simple idea that keeping track of the perspective adopted in a context furthers the chances for adequately interpreting Kant's terms in that surrounding.

Palmquist also reconstructs the Kantian system in terms of the distinctions between the various faculties. The standpoints then correspond to the interests of these faculties: theory for the understanding, morality for reason, and empirical judgment for the power of judgment. The four reflective perspectives are defined as ways in which faculties relate to each other.<sup>57</sup> As it especially is

---

However, this separation seems to be dictated by the scheme rather than by independent argument. See Westphal 1998, 338n.

<sup>57</sup> In this formulation Palmquist's mapping of Kant's system comes close to the one given in Deleuze 1984 [1963] in terms of relations between faculties. One difference is that Deleuze sees the imagination (*Einbildungskraft*) as the basic faculty in the third *Critique* (CJ), whereas Palmquist takes the power of judgment (*Urteilskraft*) to have that role. Though the imagination is of utmost importance in the aesthetic part of CJ, it has a less prominent role in the second, teleological part, which might speak in favour of Palmquist's

Palmquist's stress of the autonomy of the practical standpoint that is of value for present purposes, I will not go into the complex relations of his four reflective perspectives, though it should be noted that his distinction between empirical and transcendental perspectives roughly corresponds to the contrast between empirical and transcendental that has been applied above.

#### 4. CRITIQUE OF METAPHYSICAL INTERPRETATIONS

In this section some recent examples of metaphysical interpretations of the thing in itself in Kant are examined, with the aim of showing that some of the results achieved are less plausible than they might at first appear. The examples chosen are Langton 1998, which is discussed in some detail, Langsam 2000 (with regard to Kant's theory of truth), and Van Cleve 1999 (on principles applicable to things in themselves).

Langton proposes an interpretation according to which things in themselves and appearances are two aspects of the same thing. The crucial difference between her approach and a methodological two-aspect view is that on her account, the distinction is straightforwardly metaphysical, distinguishing two sets of properties of things, the intrinsic and the relational ones. Far from having anything to do with different epistemic perspectives, the distinction is based on the nature of things. "Things in themselves are substances qua bearers of intrinsic properties."<sup>58</sup> Phenomena are the relational properties of these substances. So there is no reference in the definition of "phenomena" to subjects of cognition or forms of sensibility. Whereas most commentators on Kant see a close connection between the doctrine of the dependence of our cognition on the transcendently ideal forms of our sensibility and that of the unknowability of things in themselves, Langton views the nature of our forms of intuition as only indirectly related to the issue of unknowability. More important than the ideality of time and space is the limitation of knowledge due

---

construal. On the relation between the imagination and the power of judgment, see also Centi 2001.

<sup>58</sup> Langton 1998, 20.

### *The Thing in Itself*

to receptivity as such. As receptive, our sensibility is dependent on our being causally affected by objects. Kant's views on the ideality of the forms of intuition are therefore not essential for deriving the thesis of our ignorance of things in themselves; the fact of receptivity is enough. This feature of Langton's reconstruction of the grounds for Kant's epistemic humility makes it an example of what Karl Ameriks has called a "short argument" for transcendental idealism. Short arguments "attempt to establish transcendental idealism [...] without going through the actual long and complex steps that Kant lays out in the [...] *Critique of Pure Reason*."<sup>59</sup> Instead of detailed examinations of the forms of intuition and the structure of judgment, they focus on some very general trait of human cognition, such as, in Langton's case, receptivity. On her account, the mere fact that empirical knowledge, by its dependence on our receiving something, involves the relation of being affected, leads to the consequence that we do not know the intrinsic properties of things but only phenomena (defined as the relational properties of things). In Ameriks view, it is implausible that Kant's case for transcendental idealism should rest on such a "global" argument, considering that Kant's detailed arguments for the ideality of time and space would then lack all bearing on the question of transcendental idealism and things in themselves, whereas Kant on the contrary seems to reach the conclusion that we have no knowledge of things in themselves on the basis of the results of the *Transcendental Aesthetic*.<sup>60</sup>

But there are also more direct reasons for doubting that Langton's interpretation is adequate. One can certainly point to passages in Kant's pre-critical works that express something resembling Langton's explication of the concept of things in themselves. To take an example not mentioned by Langton, in the *Dreams of a Spirit-Seer* (1766) Kant discusses the notions of immaterial and material substance. Of an immaterial substance he says that it could be thought to occupy a space without filling it, and should not be called extended, "for only that which occupies a space when it is separated from everything and exists *for itself* on its own *is extended*." Material substances (the elements of matter) occupy a space "in virtue of the external effect they produce on other substances," but "when these

---

<sup>59</sup> Ameriks 2001, 168.

<sup>60</sup> See Ameriks 2001.

*The Thing in Itself*

substances exist separately for themselves, and when no other things are thought of as existing in connection with them, and when there are not even to be found in them things which exist externally to each other, these substances contain no space.”<sup>61</sup> Here “substance” is taken to be that which “exists for itself on its own” or “separately,” which comes close to Langton’s characterization of things in themselves as substances considered exclusively with regard to their intrinsic properties.

But as Kant’s position develops, what he says on things in themselves begins to fit less well with Langton’s description. In the *Inaugural Dissertation* (1770), where the intellect is ascribed the capacity to know things “as they are,”<sup>62</sup> it is clear that such things can stand in relations to each other, and that these relations belong to what is known in knowledge of things in themselves.<sup>63</sup> When Kant in *Section 4* of the *Dissertation* discusses the “principle of the form of the intelligible world,” the question concerns how substances are related, and it “can only be solved by the intellect.”<sup>64</sup> Since the intellect knows things as they are, knowledge of the relations of substances to each other (regardless of whether these relations amount to real physical influence, as Kant tentatively suggests, or are better explained by pre-established harmony or by occasionalism) form part of the content of metaphysical knowledge. The various metaphysical alternatives concerning the relations of substances are theories about things in themselves, and there is no restriction on such theories to the effect that they should treat only the non-relational properties of substances. What is at stake is not the distinction between intrinsic and relational properties, but rather whether things (including their relations) are considered by the intellect, or from the point of view of sensitive cognition, via time and space.<sup>65</sup> It is also

---

<sup>61</sup> Ak. 2:323–324.

<sup>62</sup> Ak. 2:392.

<sup>63</sup> This point is also made in Ameriks 2001, 181.

<sup>64</sup> Ak. 2:407. Since Kant at this stage does not distinguish the understanding from reason, but uses the one term *intellectus*, the choice in Beck 1988 of “intellect” rather than “understanding” seems better. See also Falkenstein 1995, 367n 3.

<sup>65</sup> Kant’s distinction in the *Dissertation* between the sensitive and the intellectual worlds thus amounts to what is here called a metaphysical two-aspect theory, rather than the two-world theory the title leads one to expect.

### *The Thing in Itself*

telling that in his discussions of time and space, Kant says that these are not “objective and real, neither substance nor accident nor relation.”<sup>66</sup> This indicates that something could pertain to things in themselves (which is what “objective” and “real” mean here) and *also* be relational. The ideality of time and space does not just consist in their being relational, for there may well be relations among things in themselves, knowledge of which would belong to our knowledge of things in themselves and not to our knowledge of phenomena. That space and time are designated as ideal is because they are our subjective conditions for coordinating data of sense,<sup>67</sup> rather than because of their relationality.

Langton’s construal of the distinction between thing in itself and appearance is thus not adequate to Kant’s view in the work that bridges the pre-critical and the critical stages of his philosophical development. But perhaps Kant changed his view while composing the *Critique of Pure Reason*, so that the distinction became equivalent to that between intrinsic and relational properties? It seems plausible to assume that he did not. His way of stating that space does not belong to things in themselves shows that its ideality does not simply consist in its relationality.

Space represents no property at all of any things in themselves nor any relation of them to each other, i.e., no determination of them that attaches to objects themselves and that would remain even if one were to abstract from all subjective conditions of intuition. For neither absolute nor relative determinations can be intuited prior to the existence of the things to which they pertain, thus be intuited *a priori*. (A26/B42)

---

Indeed, Kant’s development from the *Dissertation* to the critical philosophy of CPR might be seen as the transition from a metaphysical to a methodological two-aspect view.

<sup>66</sup> Ak. 2:404; see also 2:400 (translation from Beck 1988).

<sup>67</sup> Klemme (1996, 57–59) points out that Kant conceives of this coordination in different ways in the *Dissertation* and in CPR. In the *Dissertation* it is viewed as an act of connecting due to sensibility, whereas in CPR it is merely a receptivity. See also Kemp Smith 2003 [1923], 92, and Falkenstein 1995, 93 and 96.



What restricts our cognition to appearances is thus not its relationality but its *a priori* connection to our forms of intuition. If we could “abstract from all subjective conditions of intuition” and yet have knowledge, this would be not only about “absolute” determinations, but also about the “relative” determinations of things that on Langton’s account must count as phenomenal. This is confirmed in A42/B59: “the things that we intuit are not in themselves what we intuit them to be, nor are their relations so constituted in themselves as they appear to us.” A relation can thus both appear to us and yet be thought of as having a constitution *in itself*. This is not consistent with Langton’s construal of “in itself,” according to which the term denotes the intrinsic, non-relational properties of things, excluding the very notion of a “relation in itself.”<sup>68</sup>

Langton, however, finds several passages in CPR in which Kant seems to identify that which is in itself with the intrinsic properties of things, suggesting a return to his early thought.<sup>69</sup> One such passage is in the *Transcendental Aesthetic*, where Kant says that what is “internal to the object in itself” cannot be represented through outer sense, since nothing is given to us through outer sense except representations of relations, and “through mere relations no thing in itself is cognized” (B67). Langton also adduces passages from the *Amphiboly of the Concepts of Reflection* which discuss the notion of “the inner in things” (A277/B333). For instance, Kant says that “[s]ubstances in general must have something *inner*, which is therefore free of all outer relations” (A274/B330), and that we “have no insight whatsoever into the intrinsic nature of things” (A277/B333). But, as has been pointed out by several of Langton’s reviewers,<sup>70</sup> Kant is here summarizing Leibniz’ position in order to attack it, so what Kant has to say on the internal properties of things in themselves does not necessarily reflect his own views.<sup>71</sup> According

---

<sup>68</sup> This point is also made in Setiya 2004, 82.

<sup>69</sup> Langton sees a continuity rather than a return, since she doesn’t notice that Kant in the *Dissertation* abstains from identifying the *an sich* with the non-relational.

<sup>70</sup> O’Shea 1999, 257, Rosefeldt 2001, 267–268, Walker 2002, 141.

<sup>71</sup> The statement, quoted by Langton, that says that we have no insight into the intrinsic nature of things, is in Kant’s text presented as a complaint with a somewhat unclear sense, and not as Kant’s own opinion. See A277/B333 and Rosefeldt 2001, 268.

to Kant, Leibniz one-sidedly attempted to cognize things in a purely intellectual fashion, leaving out the role of sensibility (compare A271/B327: “Leibniz *intellectualized* the appearances”). This putative cognition of the inner of things is here taken to be possible through the intellect’s being “immediately related to things in themselves” (A271/B327). The mistake Kant diagnoses in Leibniz can, in terms of Buchdahl’s model, be identified as an attempt to “realize” the object directly in a metaphysical fashion ( $T_m$ ) without recourse to the contributions of the forms of sensibility. Sensibility is for Leibniz just a distorting factor in the cognitive process, and his preferred conception of cognition is therefore an entirely conceptual affair, a view that for Kant leads to mistaking analytical judgments for synthetical ones. Far from endorsing the Leibnizian demand for cognition of the inner of things thus conceived, Kant takes this very conception of the inner, which finds its expression in the theory of the monad, to be part of the problem. In experience there is no impossibility in gradually getting deeper into the inner of nature, though this is now conceived as only “comparatively internal,” lacking the absolute character attached to the inner as thing in itself, which is “a mere fancy” (A277–278/B333–334). This “fancy” is in Kant’s view generated by the “transcendental amphiboly” that consists in “a confusion of the pure object of the understanding with the appearance” (A270/B326). The result is that purely conceptual, analytical connections pertaining to general logic are taken to yield metaphysical cognition, whereas for Kant any contentful, synthetical claims require a transcendental logic which brings with it a consideration of the conditions for experience, thereby assigning the status of appearance to the object under investigation.

The text of B67 is more difficult to interpret. One possibility is that the argument is directed against adherents of the Leibnizian view of space, according to which space is ideal in virtue of its being a system of relations holding between objects. If space is merely such a system of relations, how can we find the non-relational objects on which the relations are thought to depend, given that all we know about objects is presented through these relations? The conclusion would then be that space is ideal (as Kant and the Leibnizian agree) because it is a form of intuition, not (as the Leibnizian thinks) because its relations have objects as relata. Since objects are given as spatial, they are on the same footing as space and should be accorded

the same status, namely ideality. If the objects that are supposed to uphold the relations are nonsensible monads rather than empirical objects, Kant's point would be that we have no empirical knowledge of such things, but merely cognize ordinary objects in space. In short, "through mere relations no thing in itself is cognized" (B67).

Historically it seems plausible to connect B66–69, which was added to the *Transcendental Aesthetic* in the B-edition of the *Critique of Pure Reason*, to the objections raised against some of Kant's views by Hermann Andreas Pistorius. Specifically, Pistorius was the first critic to formulate the "neglected alternative" objection, according to which space and time may be *a priori* conditions of sensibility, and yet *also* be properties of things in themselves, a possibility not considered in the A-edition of CPR. In addition, he had pointed to difficulties in Kant's theory of inner sense with regard to its relation to apperception.<sup>72</sup>

In a review from 1786, Pistorius states that space and time are:

relational concepts that are not merely grounded in the nature of our sensibility, and that do not merely constitute its subjective forms, as Mr. K. argues, but that would also have to be considered as grounded in the nature of the *things in themselves* that appear in space and time. In this way a roughly Leibnizian concept of space and time would emerge.<sup>73</sup>

As the words "not merely" and "also" show, Pistorius doesn't deny that space and time are forms of intuition, his claim is only that this is consistent with their being grounded in things in themselves. Pistorius also argues that we obtain information about things in themselves via space and time, because they are the "pathway along which the communication between the two worlds travels,"<sup>74</sup> and that

---

<sup>72</sup> On the general importance of Pistorius' critique, see Erdmann 1878, 105–107, Beiser 1987, 188–192 and 357, and Sassen 2000, 14–16.

<sup>73</sup> Pistorius 2000 [1786], 94. This review was published in the first issue for 1786 of the *Allgemeine deutsche Bibliothek*. According to Erdmann (1878, 4), Kant started working on the emendations for the B-edition of CPR in April 1786.

<sup>74</sup> Pistorius 2000 [1786], 101. This nicely shows that the neglected alternative argument presupposes the two-world model.

### *The Thing in Itself*

this view offers a more satisfying account of ourselves as thinking subjects than the Kantian theory of apperception and inner sense.<sup>75</sup>

B66–69 fits very well within this context, as the passage begins with an investigation of the relationality of space as form of intuition (a premiss accepted by Pistorius), aiming to show that this does not allow the connection to things in themselves that Pistorius had proposed, and then turns into a discussion of inner sense and its relation to apperception. If the passage constitutes Kant’s answer to Pistorius, then his argument is built on the presupposition that space and time are *a priori* forms of intuition.

But irrespective of the exact point of Kant’s claim that we have knowledge through outer sense only of relations and therefore not of what is internal to the object in itself, it seems reasonable to assume that this comment should be taken as consonant with the general conclusion of the *Transcendental Aesthetic* that space and time are ideal, rather than as a fundamental and independent argument for our ignorance of things in themselves.

It is certainly true that the inner of things is sometimes designated by Kant as thing in itself; this does not imply that everything said about things in themselves is about the inner of things. “The inner of things” is an example of a noumenal conception, the thought of the object as metaphysically realized, closely connected to the concept of “the simple.”<sup>76</sup> For Kant the term “in itself” signals that abstraction is made from the conditions of sensibility in conceiving the thing.<sup>77</sup>

Langton helpfully locates the origin of Kant’s distinction between thing in itself and appearance in the rationalist distinction between thing and phenomenon, where the phenomena are the relational properties of things. But the traditional equation of things considered in themselves with the non-relational properties of substances is certainly not all there is to Kant’s distinction. Kant’s alteration of the conception amounts to specifying that the relations abstracted from when we consider things as they are in themselves are those

---

<sup>75</sup> Pistorius 2000 [1786], 100.

<sup>76</sup> See A283/B339 and Ak. 8:201.

<sup>77</sup> Compare Ak. 8:209: “if reason thinks a composite of substances as thing in itself (without relating it to the special character of our senses).” The parenthesized phrase indicates what one abstracts from in thinking something as thing in itself.

### *The Thing in Itself*

pertaining to our *a priori* forms of intuition. Relations among things in themselves are thus left in the realm of noumena, contrary to what we should expect if he had stuck to the traditional way of drawing the distinction. Moreover, the concept of substance is relocated to the empirical domain, and the thing in itself is left ontologically undetermined.

In an interesting exposition of Kant's account of causality, Harold Langsam finds use for the distinction between correspondence and coherence theories of truth. His discussion moves in the horizon of a metaphysical view on the thing in itself.<sup>78</sup> This interpretation guides his attempt to construe Kant's theory of truth, leading him to the claim that Kant thinks that genuine truth consists in a transcendent relation of correspondence between statements about appearances and things in themselves (albeit these are unknowable), while ordinary claims for the truth of statements about appearances, on the empirical level, are based on truth as coherence.

The idea here seems to be that a truly true statement would correspond to things as they are in themselves, but as this realm is taken by Kant to be unknowable, the need arises for an *Ersatz* truth, which will take the form of coherence. An empirical statement can then be considered true if it coheres with other empirical statements. Even though this does not satisfy the ideal requirement for truth in terms of correspondence, it is the best we can have.

Behind this interpretation we find the metaphysical view according to which things in themselves (which Langsam equates with the intrinsic properties of things) constitute the ontologically real domain, whilst appearances lack full-fledged reality in virtue of being mind-dependent properties. Expressed in perspectival terms, what Langsam does is to construe Kant as setting the realization of the noumenon in positive respect as the standard for truth, and then, seeing that we have no way to accomplish this realization, making do with a second-best option to explain the possibility of empirical truth.

According to Langsam, when the methodological two-aspect view claims that an object (for instance a drop of water) can be considered to be both a thing in itself in the empirical respect (and thus not

---

<sup>78</sup> Langsam (2000, 186) aligns himself with the metaphysical two-aspect view.

appearance, as the rainbow might be said to be), *and* appearance in the transcendental respect (hence not a thing in itself), it illegitimately assumes that the meanings of the terms “thing in itself” and “appearance” differ on the respective levels.<sup>79</sup> Now, this is not entirely clear, as one might rather say that it is a question not of different meanings, but of differing contextual applications of terms with univocal meanings (as exemplified with the word “left” in section 3 above). A clarification of the issue would presumably require a more differentiated account of meaning. At any rate, Langsam’s proposal is that Kant’s claim is that it always is false, on any level, that the drop of water is a thing in itself. It is merely appearance.<sup>80</sup> Nonetheless, Kant deems it appropriate to recognize a notion of empirical truth, since we are constrained by the nature of our cognitive faculty to take appearances as real things.<sup>81</sup> As this constraint ensures that an empirical statement purports to refer to appearance as if it were a thing in itself, the statement is false as measured with the standard of correspondence truth, but with respect to the other notion, that of empirical truth, it is true. Empirical truth is not a correspondence notion but a coherentistic notion relating appearances to each other.

Several questions are prompted by this account. In the first place, if it is illegitimate to let terms such as “thing in itself” and “appearance” have different meanings with respect to empirical and transcendental points of view, as Langsam argues, why would it be any less problematic to assign different meanings to the term “truth” on the respective levels? If univocity is important, it should be so also in this case. Secondly, what is the relation of this account to that offered by the methodological two-aspect theorist? The latter, call her the methodologist, might put it like this:

**Empirical perspective:** It is true that the drop of water is a thing in itself and not an appearance.<sup>82</sup>

---

<sup>79</sup> Langsam 2000, 174.

<sup>80</sup> Langsam 2000, 174. The rainbow example is in A45–46/B62–63.

<sup>81</sup> Langsam 2000, 177–178 (based on A191/B236).

<sup>82</sup> Note that this is the use of “thing in itself” in empirical respect discussed in the beginning of section 3.

**Transcendental perspective:** It is true that the drop of water is an appearance and not a thing in itself.

Langsam's construal of Kant's position can be put as follows:<sup>83</sup>

**Empirically true:** The drop of water is a thing in itself and not an appearance.

**Transcendentally true:** The drop of water is an appearance and not a thing in itself.

It might now seem as if the difference is merely notational. What the methodologist accomplishes by means of a perspectival distinction, Langsam gets by detaching "true" from the statement and putting it where the methodologist has "perspective." However, in view of Langsam's coherentistic construal of empirical truth, the accounts differ after all. For the methodologist has a choice: she is free to construe truth in the empirical perspective as correspondence.

Thirdly, one can question the interpretation of Kant's example in A45–46/B62–63 on which much of Langsam's account is based. Kant says that if we rest content with establishing that on the empirical level the raindrops are things in themselves and the rainbow is an appearance, we will think that the raindrops are things in themselves also in the transcendental respect. His point is that without the transcendental distinction there is no reason not to take an empirical thing to be a thing in itself also on the level of philosophical reflection, and then there will be no way to escape all the problems and antinomies that transcendental realism gets entangled in; whereas if we take the transcendental step, we get the advantages of transcendental idealism. But we do not have to take Kant to say that the transcendental status of the raindrop, its being an appearance, amounts to the truth also when the situation is viewed from the strictly empirical perspective. In fact, Langsam's attribution of this claim to Kant blurs two distinct levels, so that the transcendental meta level perspective is applied also in empirical contexts. The skeptical consequence of this is softened by the adoption of the empirical notion of truth in terms of coherence, but Langsam is clear on its second-rate status compared to the transcendent correspondence truth.

---

<sup>83</sup> Compare Langsam 2000, 179.

*The Thing in Itself*

What does Kant actually say about truth?<sup>84</sup> He points out that the nominal definition of truth is “the agreement of cognition with its object” (A58/B82; A237/B296; A820/B848). It is this agreement or correspondence that defines truth for an empirical judgment. The primary locus of truth as correspondence is not in relation to things in themselves; on the contrary it is the correspondence of empirical statements to objects of experience that is relevant. But in order to explain the possibility of true judgments with respect to appearance, the whole transcendental theory of empirical cognition has to be brought in. The transcendental conditions for experience are necessary conditions for the possibility of empirical judgments (true or false). In this way, Kant attempts to provide a meta theory on how truth as correspondence is possible. It is in this light that one should read passages that put truth in the vicinity of coherence.<sup>85</sup> They point out that nature, to which true judgments correspond, is itself a coherent system of experience under rules. Rather than being the definition of truth, coherence characterizes the ensemble of true empirical statements, since these refer to experience which is a coherent system.<sup>86</sup> Coherence is important for epistemic justification, serving as indicator of veridical cognition.<sup>87</sup> Yet truth as such is by definition correspondence, and in the case of empirical cognition what judgments correspond to are objects of experience as based on intuition.

Whether or not there is an element of coherence entwined in Kant’s correspondence view of truth, there is no basis for a partitioning that places correspondence in relation to noumena and

---

<sup>84</sup> The following sketch owes much to the fine exposition in Nenon 1986.

<sup>85</sup> For instance, in A451/B479 nature as “the connection of appearances necessarily determining one another in accordance with universal laws” is aligned to “the mark of empirical truth”, and in *Prolegomena* (Ak. 4:291) the connecting of intuitions “according to rules for the combination [*Regeln des Zusammenhanges*] of all cognition in one experience” is specified as what makes truth (but also illusion) possible.

<sup>86</sup> See for instance Ak. 4:455: “experience – that is, coherent knowledge of sensible objects in accordance with universal laws” (translation from Paton 1991).

<sup>87</sup> See also A820/B848 on intersubjective agreement as a consequence of truth, and Nenon (1986, 241–243) on this passage; cf. Nenon (1986, 230–231) on the role of coherence.



coherence in the phenomenal realm. Langsam's construal of Kant's theory locates correspondence on the wrong level, because it is conditioned by a metaphysical view of things in themselves which assumes the priority of an unreachable realm, rather than, as in the methodological view, considering this realm as a thought object problematically and abstractly conceived.

A third example of how a metaphysical approach (this time of the two-world type) can lead one astray is provided by Van Cleve's discussion of a question connected to Kant's critique of Leibniz. In the *Amphiboly* chapter, Kant seems to grant that some Leibnizian principles, not applicable to appearances, nevertheless are valid for things in themselves. Among these are the principles that complexes are composed of simples, that relations must have something inner as ground, and that indiscernible objects are identical (A270–289/B326–346). Van Cleve's point is that the two-world theorist can explain this simply by saying that such a principle holds for one kind of object (things in themselves) but not for the other kind (appearances). A two-aspect theorist, on the other hand, must say that the principle is valid for things considered in abstraction from sensibility though not valid for these things as they appear.<sup>88</sup> But this means that objects would consist of simples when considered in one way, but strangely enough have another mereological structure when considered in another way, or that two objects are identical if considered in themselves, but not identical if considered in relation to our sensibility.<sup>89</sup>

Van Cleve's reading is based on the attribution of definite ontological properties to things in themselves, which are taken to be accessible to the metaphysician's inspection, so that it can be known that they have a certain structure, such as simplicity. In its context, however, Kant's claim is that these are principles which Leibniz must hold because he attempts to attain knowledge about things "*through mere concepts*" (A284/B340). It is not his own ontological doctrines on things in themselves that are presented here, it is presuppositions of the Leibnizian attempt to reach noumenal cognition solely by the intellect.

---

<sup>88</sup> Van Cleve 1999, 149.

<sup>89</sup> Van Cleve 1999, 149–150.

Still, there is something here in need of explanation. Why does Kant say that the simple is the “foundation of the inner of things in themselves” (A274/B330)? How can he claim that for things in themselves, the parts are never conditioned by the whole (whereas the reverse holds with respect to space, which therefore is appearance)?<sup>90</sup> As Kant rather frequently states how we must think that noumena are constituted,<sup>91</sup> it is understandable that commentators have concluded that a certain amount of knowledge, or at least of essentially correct thinking, about abstract properties of things in themselves is after all allowed in the critical philosophy.<sup>92</sup> Kemp Smith points out how close this seems to be to the position of the *Dissertation* that allows pure thought to determine the nature of things as they are.<sup>93</sup>

The key to resolving this difficulty is to take seriously Kant’s claim that these structural properties of things in themselves represent the way we have to think when we abstract from the conditions of sensibility. We “distinguish the object represented in general and *in abstracto* from [our] way of intuiting it; thus there remains to us a way of determining it merely through thinking that is, to be sure, a merely logical form without content, but that nevertheless seems to us to be a way in which the object exists in itself (*noumenon*), without regard to the intuition to which our sensibility is limited” (A289/B346).

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. Concepts without intuitions are empty. As no object is introduced by mere concepts, some kind of intuition would be required in order to reach cognition. For this to be different from ordinary cognition of appearances, the intuition would have to be intellectual. If we possessed such intuition, we would cognize the noumenon in positive respect. But the attempt to cognize the object purely conceptually, without any intuition at all, which basically is what Kant in the *Amphiboly* ascribes to Leibniz, is doomed merely to reproduce the structural features of concepts.

---

<sup>90</sup> *Prolegomena*, Ak. 4:286.

<sup>91</sup> For another example, see Ak. 4:507.

<sup>92</sup> Compare Adams 1997, 810–811.

<sup>93</sup> Kemp Smith 2003 [1923], 419.

The particular form that such intuitionless thinking takes must reflect the nature of concepts. We can get some information about Kant's views on this by looking into his theory of concepts as it is expounded in the context of the distinction between synthetic and analytic judgments (B10–13). The distinction is based on a theory of conceptual containment. If the conceptual content of the predicate is contained in the concept of the subject, the judgment is analytic. For instance, the proposition "Gold is a yellow metal" is analytic, because the concepts "yellow" and "metal" are contained in the concept "gold."<sup>94</sup> A concept is usually complex, and to analyze it amounts to finding its content, the concepts by which it is composed. The constituents can in their turn be further analyzed. But, as the *Prize Essay* (1764) explains, though the analysis can continue, it has to reach an end sooner or later: there must be elementary concepts, whether we succeed in pushing the analysis that far or not.<sup>95</sup> For a human being to be able to think it, the concept has to consist of a finite number of subconcepts.

Kant's theory of concepts, which (except for innovations such as the analytic-synthetic distinction) is to a large extent in line with what most German thinkers from Leibniz on accepted,<sup>96</sup> offers a possibility of explaining the particular claims about how we have to think of things in themselves, claims that otherwise seem entirely arbitrary. That things in themselves must be thought of as absolutely simple reflects the simple, elementary concepts with which conceptual analysis terminates. That we have to think of things in themselves on the model of wholes composed of parts, the parts of which are not conditioned by the whole (contrary to how it is with space), also finds its explanation in the theory of concepts. Concepts depend on their subconcepts, but the subconcepts contained in a concept are not dependent on the concept. From "gold" we can infer the subconcept "metal", though from "metal" we can't infer "gold".<sup>97</sup> Thus "metal"

---

<sup>94</sup> See *Prolegomena*, Ak. 4:267.

<sup>95</sup> See Ak. 2:280: "it is obvious from the start that the analysis will inevitably lead to concepts which are unanalysable."

<sup>96</sup> See De Jong (1995, 620–627) on this tradition, and Parkinson (1995, 201) on Leibniz.

<sup>97</sup> Compare Ak. 2:202: "the consequence is really identical to part of the concept [*einem Teilbegriffe*] of the ground".

does not contain “gold” (although “gold” as regards extension *falls under* “metal”), and in that sense the subconcept is not dependent on the composite concept.

According to Kant’s diagnosis of Leibniz’ intellectualistic mistake, it is such conceptual relations that are falsely conceived of as pertaining to objects in themselves, so that the structure of conceptual containment is misrepresented as a corresponding substantial property of the noumenon. The same goes for the principle of the identity of indiscernibles. Two concepts which are indiscernible are the same, and this is projected on the metaphysical plane as a substantial principle. But for the cognition we in fact have, which is not solely conceptual but also requires sensibility, the principle loses its force: two objects that are indiscernible as to their conceptual determinations can yet be non-identical, since they can have different spatio-temporal locations.<sup>98</sup>

The kind of projection of conceptual relations on objects of thought delineated here is not the sole type of noumenal thinking that Kant examines. A most important attempt to get from concept to object in itself is by way of the transcendental ideas, which Kant treats in the *Dialectic*. In this case too, it is a question of conceptual relations being hypostasized. The three forms of inference of reason, represented in the categorical, the hypothetical and the disjunctive syllogism (which three forms Kant considers to exhaust the logical theory of inference), generate a dialectical illusion which gives us the thought of noumenal objects attainable by pure reason.<sup>99</sup> These are the ideas of the soul, the world and God, as conceived of in “special metaphysics.” I shall not broach the difficult issue of the derivation of ideas from the logical operations of reason, and neither shall I examine the *Paralogism*, *Antinomy* and *Ideal* connected to the three ideas. The ideas point to objects that are illusory from the theoretical standpoint, but Kant nonetheless finds it significant to conceive of such noumena from the practical standpoint. This involves a shift of perspective of the kind thematized by Palmquist. In the concluding section this will be briefly looked into.

---

<sup>98</sup> Kant’s view depends on his claim that spatial orientation is not possible to express in a purely conceptual way, an issue that can’t be pursued here.

<sup>99</sup> See Grier 2001 for a book-length treatment of the theme of transcendental illusion.

## 5. CONCLUDING REMARKS: ETHICAL DIMENSIONS

Is the two-aspect approach adequate for understanding Kant's practical philosophy? In the preface to the *Critique of Practical Reason*, Kant claims that freedom is the only one of the ideas of reason "the possibility of which we know [*wissen*] *a priori*," and we know it because we know that we are bound by the moral law.<sup>100</sup> It might appear as if this involves an addition to our theoretical cognition, in that freedom, the actuality of which could not be established theoretically, is now known via the moral law. But rather than an addition to theoretical cognition, it is an addition of a new perspective beside that of theoretical cognition. In the theoretical perspective, no question of morality arises; only how things are is under consideration. To enter the domain of morality, that is, to switch from the theoretical to the practical perspective, involves a transition from *is* to *ought*.<sup>101</sup> Kant claims that the moral law is present for each person as a consciousness of obligation. This is not to be found in the scientific investigation of nature which is the business of the theoretical perspective.

In order to make it possible to assign the knowing of freedom to the practical perspective, there has to be a way of construing *Wissen* in this domain that keeps it distinct from theoretically attained cognition. For this purpose it is useful to begin with a look at the characterization of *Wissen* in CPR. In A822/B850, Kant distinguishes three "stages" in taking something to be true (*Fürwahrhalten*): having an opinion, believing, and knowing (*Wissen*). In knowing, there is both subjective and objective sufficiency in the reasons for taking something to be true. Objective sufficiency is also called certainty (*Gewissheit*). In regard to pure reason, where no empirical confirmation is possible, there is no point in having an opinion. In such matters of reason as pure mathematics and the principles of morality, "one must know, or else refrain from all judgment" (A823/B851). This shows that there is conceptual room for a practical knowing. Moral principles are counted to what can be known, though they pertain to what ought to be and not to what is. A little later, Kant

---

<sup>100</sup> Ak. 5:4.

<sup>101</sup> Compare for instance A802/B831.

introduces a distinction between logical and moral certainty, corresponding to the theoretical and the practical perspective respectively (A829/B857). Moral certainty is here said to be “interwoven with my moral disposition” (*Gesinnung*).

Now, as knowing implies certainty, the knowing of moral principles that Kant has made room for must be an instance of moral certainty. There is thus a place in the conceptual landscape of the first *Critique* for a practical knowing, based on the complex moral concept of *Gesinnung*.<sup>102</sup> It is this kind of knowing of the moral law that is the topic of the *Critique of Practical Reason*. This knowing brings with it a knowing of freedom, since freedom is the condition of the possibility of the moral law. As the knowing of freedom is derived from the knowing of moral obligation,<sup>103</sup> it too is a practical certainty, interwoven with moral *Gesinnung*. Kant’s use of *Wissen* in his practical philosophy can thus be accommodated in a perspectival interpretation.<sup>104</sup> For instance, it fits with Palmquist’s conception of a separate practical standpoint. Likewise, it goes well with Buchdahl’s model. In Buchdahl’s terms, the attainment of knowledge in moral philosophy is a practical realization of the object as positive noumenon, which is distinct from a metaphysically conceived positive noumenon, in that it is not based on theoretical speculation but on phenomena of a different order, such as moral appraisals and the consciousness of duty.<sup>105</sup>

Keeping the theoretical and the practical perspectives or standpoints apart also helps explaining the seemingly incongruous use of *Wissen* and *Glauben* in the Preface to the second edition of CPR, where Kant famously says that he “had to deny *knowledge* in order to make room for *faith*” (Bxxx). As freedom is clearly included in that of which knowledge is denied (it is mentioned immediately above as a necessary assumption for the practical use of reason), this seems to be a direct contradiction of the claim in the second *Critique* that we know that we are free. As the two texts belong to about the

---

<sup>102</sup> For an account of *Gesinnung*, see Allison 1990, 136–145.

<sup>103</sup> Though freedom is the “*ratio essendi* of the moral law, the moral law is the *ratio cognoscendi* of freedom” (Ak. 5:4).

<sup>104</sup> Compare Steigleder 2002, 100: “Dieses Wissen [...] verbleibt ganz auf der Ebene des Praktischen”.

<sup>105</sup> See Buchdahl 1992, 323–325.

same period of composition, the occurrence of such a contradiction is hard to believe. On the present interpretation, the contrast between knowledge and belief in Bxxx is drawn from the point of view of the theoretical domain. What is not theoretical knowledge is considered as belief or faith. From the practical perspective, on the other hand, the consciousness of moral obligation offers a criterion for practical knowledge, which is thereby distinguished from belief. What is known in this perspective does not count as knowledge in the theoretical domain. Knowing that we are free is certainly an addition to our knowledge, generally speaking, but then in terms of existential rather than theoretical knowledge.<sup>106</sup> Kant speaks in the *Critique of Practical Reason* of the “objective reality of a pure concept of the understanding in the field of the supersensible” which is “of only practical applicability and has not the least influence on theoretical cognition.”<sup>107</sup> Thus, in the practical perspective the will’s causality is represented as a noumenon, the concept of which is the merely formal “thought of an object in general” when viewed theoretically, but which obtains practical signification through the moral law.<sup>108</sup>

In Kant’s view, the possibility of admitting two distinct perspectives of the theoretical and the practical, without reducing the latter to a merely fictional “as if”-sphere, presupposes that there is a solution to the antinomy between freedom and causal determination. This opens large questions which will not be entered here. But it should be noted that Kant’s solution (according to which the human being in one respect, considered as a natural object, is causally determined and in another respect, considered as noumenon, may be free), explicitly relies on a two-aspect model. It is crucial for Kant’s solution that it is the same human being that in one respect is a part of nature and in another respect adopts the standpoint of freedom.<sup>109</sup>

It should also be noted that the perspectival approach does not preclude the metaphysical possibility that there may be more to reality than what is attainable in theoretical cognition. Kant stresses

---

<sup>106</sup> Accordingly, Kant suggests that the proper expression is not “*it is* morally certain” but “*I am* morally certain” (A829/B857). (I assume that this can be applied not only in the theological context in which it occurs.)

<sup>107</sup> Ak. 5:56.

<sup>108</sup> Ak. 5:50.

<sup>109</sup> See for instance A540–541/B568–569.

that it must at least be logically possible that the freedom manifested in the practical perspective co-exists with the theoretically known causal determination of nature. If the idea of freedom were to lead to a logical contradiction, there would be no possibility to reasonably entertain it, not even in the guise of an “as if”-freedom.<sup>110</sup> But even though transcendent metaphysical possibilities should not be dogmatically ruled out, they can hardly serve as the starting point for interpreting Kant, as in the two-world interpretation and the metaphysical version of the two-aspect view.

It is sometimes said that in his personal beliefs Kant retained the metaphysical ideas characteristic of the Leibniz-Wolffian tradition.<sup>111</sup> As indicative of this, Westphal (1997, 241) points to the fact that in his critical period Kant continued to teach traditional school metaphysics, using Baumgarten’s *Metaphysica* even when a textbook accurately presenting the critical philosophy was available. Now, one has to ask why, if he thought that Baumgarten’s metaphysics was correct, Kant didn’t say so in his own philosophical works? And if that is just what he did, though in a somewhat ineffective way (which perhaps is what some of the metaphysically oriented interpretations try to show), then why could he not teach his own philosophy, as it then would have harmonized with his personal beliefs?

As for his use of Baumgarten’s book in teaching, Kant did not refrain from pointing out faults in it, and he also lectured on quite a few of the tenets of transcendental philosophy, as is readily seen for instance in the *Metaphysik Mrongovius* (1782–1783).<sup>112</sup> The reason why Kant continued to use Baumgarten’s book might well be that he thought it difficult for students to understand transcendental philosophy without a knowledge of school philosophy. There is, however, another possible explanation for Kant’s conservative teaching policy (disregarding indications in student lecture notes

---

<sup>110</sup> Compare Ak. 4:456: “if even the thought of freedom contradicts itself or contradicts nature [...], it would have to be given up altogether.”

<sup>111</sup> Baertschi (2004, 192 and 207–208) shows what prominent role appeals to such alleged personal beliefs played for the German metaphysical school of Kant interpretation associated with M. Wundt and H. Heimsoeth among others.

<sup>112</sup> Ak. 29:747–940; Kant 1997b.



speaking against conservativeness). In his essay *An Answer to the Question: What is Enlightenment?* (1784), Kant famously distinguishes public and private uses of reason.<sup>113</sup> What is at first bewildering for the modern reader is that Kant here uses the terms “public” and “private” contrary to their current meanings. The use made of reason in a particular civil post or office is private. A clergyman, for example, must express the views of the church which he represents. If, as a scholar, he expresses an opinion divergent from that of his church, nothing can be said against it, as long as he does this in respect of being a scholar who addresses the entire reading public in a free debate. This latter is the public use of reason. As a representative of an organization a person’s reasoning remains private, not in the modern sense of expressing individual opinion, but in the sense of being restricted.<sup>114</sup> The official representative is deprived of his free, public use of reason, which can only take place outside of the organization to which he is bound. One might surmise that the same goes for university teachers, which would cast some light on the circumstances described by Westphal. As a university professor, Kant may have considered himself bound to his official – private – role, which consisted in teaching doctrines conformable to the university’s. As there were competing approaches to philosophy, it is not easy to say what the official line would have been: Wolffianism, Lockean empiricism, or perhaps even pietism; at any rate, it was definitely not transcendental idealism. As a free writer and scholar, on the other hand, Kant could disseminate his critical philosophy, a philosophy which from the very beginning was conceived of as a defence of reason in its public function.<sup>115</sup>

---

<sup>113</sup> See Ak. 8:37–38.

<sup>114</sup> O’Neill (1992, 298) connects “private” to *privatus* in the sense of “deprived.”

<sup>115</sup> See O’Neill 1992 on this theme.

# Kant's Practical Deduction of Moral Obligation in *Groundwork* III

The third part of Kant's *Groundwork of the Metaphysic of Morals*<sup>1</sup> is notorious for its obscurity. The unclarity pertains not only to the course of the argument delivered in the text, but also to its very purpose. What is it that Kant gives a deduction of – the categorical imperative (see for instance G 461), or freedom (as suggested in G 447: “nor can we as yet make intelligible the deduction of the concept of freedom”)? Further puzzlement is introduced by Kant's later claims in the *Critique of Practical Reason* that the moral law is a “fact of reason,” and that no deduction can be given thereof, claims that seem to contradict the G III account.

According to the influential interpretations of Karl Ameriks and Henry Allison,<sup>2</sup> G III is an attempt to justify the moral law from non-moral premises. This involves giving a theoretical proof of freedom (in the transcendental, non-compatibilist sense), and then using this result as a means for establishing the validity of the moral law. The step from freedom to morality would be licensed by the reciprocity between these concepts.<sup>3</sup> The argument is taken to be intended as a theoretical one: Kant does not merely establish our freedom in the practical perspective.

Ameriks links the deduction in *Groundwork* to a rationalistic belief in the *cogito* as a source of knowledge of our freedom, a view

---

<sup>1</sup> In referring to Kant's *Groundwork*, I follow the translation in Paton 1991, with but minimal alterations. “G” refers to *Groundwork* and page numbers to the pagination of the *Akademie* edition (Vol. 4).

<sup>2</sup> Ameriks 1982, 189–233; Allison 1990, 214–229.

<sup>3</sup> See the discussion of this in Allison 1990, 201–213.

“that Kant had held for some time and simply had not gotten around to submitting to a thorough critique.”<sup>4</sup>

In Allison's reading, Kant proceeds somewhat more cautiously. He does not argue directly from reason's theoretical spontaneity to a corresponding freedom of the will. Rather, he tries to establish our membership in an intelligible order above the world of sense. This step would guarantee that our consciousness of freedom is not illusory, and so a deduction of morality can follow. But, as Allison shows, our membership in the intelligible world can not be proven, and so Kant's entire deduction merely begs the question.<sup>5</sup>

The question must be posed, however, whether Kant's argument should not rather be considered to move mainly in the practical perspective. Some of the questionable passages can be read as descriptions of intellectual operations performed in acting, and not as steps in a theoretical proof with purely non-moral premises. A practical deduction (by which I mean a justification that takes place from the practical point of view of action) can involve non-moral premises without constituting a theoretical proof (that is, a proof pertaining to the theoretical perspective of knowledge). There is to be sure a theoretical argument annexed to the deduction, but on this view its function is only to prove the logical possibility (non-contradictoriness) of freedom, a point that is essentially that of the Third Antinomy of the *Critique of Pure Reason*.

Kant's exposition in G III will be briefly studied from this point of view, leaving other issues in this rich text untouched.

## 1. NEGATIVE AND POSITIVE FREEDOM

At the end of part II of the *Groundwork*, Kant indicates the need to prove that the moral principle is “no mere phantom of the brain.” To do this amounts to showing that it is a principle *a priori* (G 445). What is required is a deduction (that is, a justification) of the categorical imperative as an *a priori* condition of action, constituting obligation.

---

<sup>4</sup> Ameriks 1982, 214.

<sup>5</sup> Allison 1990, 228.

## *Kant's Practical Deduction of Moral Obligation*

G III begins with a consideration of the concept of freedom. Regarded as independence of the causality of natural law, it is defined negatively. The positive content of the concept is not indeterminacy or randomness, but another law to which a causality corresponds, a law given by the will to itself. This is autonomy, contrasting to the heteronomy of natural necessity. Kant connects this to his earlier analysis of the concept of the moral law as a self-imposed legislation, and so freedom of the will is the same thing as a will under moral law (G 446–447). A route from freedom to morality is thus opened by the analytical connection between these concepts, but Kant makes clear that there is much to do before it can be taken with regard to the human case. The analytical connection does not warrant the categorical imperative in so far as this is a synthetic proposition. So even if our freedom could be established, something would be missing in an account that proceeded analytically.

## 2. ACTING UNDER THE IDEA OF FREEDOM

If freedom is to be demonstrated, this must be done *a priori* (G 448). Since freedom is an idea of reason, no empirical state of affairs can ever confirm or disconfirm it. Morality concerns us as rational beings, and so freedom must also apply to all rational beings (to the extent that they are endowed with a will). We are not looking for a property particular for human beings, but for a property of the will of any rational being.

At this point in his exposition, Kant asserts that “every being who cannot act except *under the idea of freedom* is by this alone – from a practical point of view – really free” (G 448). He adds that practical reason, i.e. the will of a rational being, must regard itself as free. It is also pointed out that this being so, we are not required to prove freedom from a theoretical point of view: “In this way we can relieve ourselves of the burden which weighs upon theory” (G 448n). Whatever the merit of Kant’s view here, according to which a will that must presuppose its own freedom is as bound by the laws of freedom as one that actually *is* free, this note indicates strongly that the whole deduction is conducted in the practical perspective rather

than in the theoretical. But the issue is not clear-cut, since we are still only at the beginning of G III (maybe the subsequent problem of the circle will force us to transfer to the theoretical point of view), and also since Kant's claims here are stated in the first person: "Now I assert," "And I maintain" (G 448). This is presumably a signal that what is said is not considered to be sufficiently established at this point, leaving us uncertain as to its force.<sup>6</sup>

### 3. THE CIRCLE AND THE TWO STANDPOINTS

Kant's discussion now turns to the question of how the moral law can be binding for us, why we subject ourselves to the moral "ought." This might seem a somewhat confusing introduction of a new problem. Why would it be relevant for our deduction of morality from freedom? The answer appears to be that if we derive the moral law from freedom, we have not said enough about the fact that we are not purely rational beings, but also "affected by sensibility."<sup>7</sup> For us the moral law takes the form of an imperative, our will is not "holy" (cf. G 439). The distinction here is between the moral law as the (logically unbreakable) law for the actions of an ideally rational being, and the categorical imperative as the way the law is given for only partly rational beings like us, capable of breaking the law but obligated not to.<sup>8</sup> These two aspects of the moral law are what Kant describes as an analytical and a synthetical connection respectively.

The dilemma this poses is that either we have a will and consider ourselves as free, in which case the moral law is valid for us (assuming the reciprocity of freedom and morality), but then we are holy wills to which the concept of obligation does not apply (compare G 439); or we are heteronomously conditioned by natural causes, in which case morality is out of the question for us. We must be able to justify the moral law as valid for us, without abolishing the very question of morality by defining ourselves as holy wills. We must

---

<sup>6</sup> Compare O'Neill 1989, 54–55.

<sup>7</sup> The importance of this theme in G III is rightly stressed in Schönecker 1997 and 1999.

<sup>8</sup> Compare Allison 1990, 278n 1.

*Kant's Practical Deduction of Moral Obligation*

also give an account of how the moral law (considered as the law of a holy will) can take the form of a categorical imperative (the ought that presents itself to us).

This problem concerning the possibility of moral obligation is an important contextual background to Kant's famous circle, and its solution is also, I think, the main task of the deduction. The threat of circularity arises because of the analytical connection between freedom and morality:

In this, we must frankly admit, there is shown a kind of circle, from which, as it seems, there is no way of escape. In the order of efficient causes we take ourselves to be free so that we may conceive ourselves to be under moral laws in the order of ends; and we then proceed to think of ourselves as subject to moral laws on the ground that we have described our will as free. Freedom and the will's enactment of its own laws are indeed both autonomy – and therefore are reciprocal concepts – but precisely for this reason one of them cannot be used to explain the other or to furnish its ground. (G 450)

We assume freedom just because we want to reach morality as our conclusion, and then we happily justify our putative subjection to the moral law by pointing to our freedom. But we haven't explained or understood anything.<sup>9</sup> At this juncture, there is a very natural urge to ask for a theoretical proof of freedom, in order to get a non-question begging foundation for morality. *Prima facie*, this is what Kant offers. He says that there still remains one way out, namely, to make a distinction between two standpoints. One standpoint is taken when we assume that we are “*a priori* efficient causes,” and a different one when we consider ourselves in terms of our actions as natural effects that just happen in nature (G 450). This is of course the doctrine of transcendental idealism, with its distinction between things in themselves and appearances. But it is depicted as a distinction that is drawn by the most common understanding, “without any need for subtle reflexion”(G 450). It is noteworthy that the distinction is said to arise in ordinary reflexion. The two standpoints – the sensible

---

<sup>9</sup> As argued by Schönecker (1997 and 1999, ch. 5.2), the circle in question takes the form of a *petitio principii* rather than a *circulus in probando*.

world as passively observed and the intelligible world (*Verstandeswelt*) as linked to activity – are, so to speak, phenomenologically accessible to anyone. That counts also for our thinking of ourselves as members of the intelligible world, in virtue of our finding reason in ourselves (G 452).

“Presumably, the point here is that our membership in the intelligible world provides the needed nonmoral premise from which our freedom and [...] our subjection to the moral law can be derived,” says Allison.<sup>10</sup> He points out the severe difficulties in going from the spontaneity manifested in reason to membership in the intelligible world and therefrom to the possession of a will. According to Allison’s analysis, Kant slides from the thought of the intelligible world (*Verstandeswelt*) as whatever is nonsensible (the noumenon in negative respect) to the intelligible world (*intelligibelen Welt*) as governed by the moral law (the noumenon in positive respect).<sup>11</sup> Only by this step is Kant able to escape the circle. But the cost is very high, since it amounts to that kind of unjustified jump into the transcendent sphere otherwise condemned by the critique of reason.

However, if the argument is construed as conducted in the practical perspective, its invalidity is less obvious. As described above, the problem of the circle is connected to the requirement for two standpoints: one heteronomous and one autonomous. Our main difficulty, then, is not that we cannot theoretically prove transcendental freedom, but that we, *from inside the practical sphere*, can’t see how we can be both free and unfree (autonomous and heteronomous). Kant’s strategy is to appeal to the two standpoints, as perspectives arising in reflexion. This does not amount to a proof of anything, but it opens a space for the thought of freedom without abrogating the fact that the moral law presents itself to us as an imperative (which it could not do for a purely rational being).

We see now that when we think of ourselves as free, we transfer ourselves into the intelligible world as members and recognise the autonomy of the will together with its consequence – morality; whereas when we think of ourselves as under obligation, we look

---

<sup>10</sup> Allison 1990, 221.

<sup>11</sup> Allison 1990, 227.

*Kant's Practical Deduction of Moral Obligation*

upon ourselves as belonging to the sensible world and yet to the intelligible world at the same time. (G 453)

In a passage just above this one, Kant removes the reservation indicated by his use of the first person at G 448, concerning our having to view ourselves under the idea of freedom. The introduction of the concept of the intelligible world provides the justification for what was earlier asserted as a mere opinion:

As a rational being, and consequently as belonging to the intelligible world, man can never conceive the causality of his own will except under the idea of freedom; for to be independent of determination by causes in the sensible world (and this is what reason must always attribute to itself) is to be free. (G 452)

According to the interpretation of the deduction as a theoretical proof, we start from reason's spontaneity and reach the intelligible world, where we find the free will needed for deriving morality.<sup>12</sup> The passage just quoted is admittedly quite possible to read in that way. But on the present reading, the independence that reason attributes to itself is not theoretical in the sense of being a provable truth from a cognitive point of view. It is, rather, the stance of independence taken for granted in our thinking and acting. Reason's spontaneity involves the use of ideas (compare G 452), among them the idea of freedom. This idea is connected to the concept of the intelligible world, which coupled to the concept of the sensible world provides the conceptual space needed for morality in the imperative form it has for us. Thus, we find Kant saying that "the idea of freedom makes me a member of an intelligible world" (G 454); far from finding my freedom in the intelligible world (as in Allison's rendering of the deduction), I become a member of it by viewing myself under the idea of freedom.<sup>13</sup> Viewing myself under the idea of freedom is the stance taken up in thinking and acting, when I consider myself an agent (cf. G 448). "The concept of the intelligible world is thus only *a point of view* which reason finds itself constrained to adopt outside appearances *in order to conceive itself as practical*" (G 458).

---

<sup>12</sup> Compare Allison 1990, 223–224.

<sup>13</sup> This point is also made in McCarthy 1985, 31–32.



#### 4. THE THEORETICAL DIMENSION

So far, we have construed Kant's argument as practical. But there is also a theoretical argument entwined in his discussion. It begins in the section *On the extreme limit of all practical philosophy*, where the task of speculative (theoretical) philosophy is specified. A dialectic of reason arises out of the apparent contradiction between freedom of the will and natural necessity. The contradiction is unavoidable if the subject conceives itself in the same sense when it takes itself as free as when it considers itself as determined by natural law (G 456). Theoretical philosophy must show that these two characteristics can be combined in the same subject considered in different respects, as thing in itself and as appearance. This corresponds to the result of the Third Antinomy of the *Critique of Pure Reason*: transcendental idealism makes room for two standpoints, and thereby shows that freedom is logically compatible with the causality of natural law, without proving that there actually is freedom.

Theoretical philosophy had to solve this problem for the sake of practical philosophy, because no conception of freedom in a practical perspective would be acceptable if it could be shown that the very idea of freedom (in the strong, incompatibilist sense) is logically excluded by the causal necessity of nature, or that it is self-contradictory (G 456).

In my view, this is the only clearly theoretical consideration in the *Groundwork* deduction. Against this, one could point to Kant's claims about spontaneity and our possession of reason as indicating an attempt to give a speculative grounding of freedom. Shortly after the assignment of a task for speculative reason that we have just seen, Kant writes the following:

The lawful title to freedom of will claimed even by ordinary human reason is grounded on a consciousness – and an accepted presupposition – that reason is independent of purely subjective determination by causes which collectively make up what belongs to sensation and comes under the general name of sensibility. (G 457)

It would seem that reason's independence of sensibility, its spontaneity, is something that we are aware of, something we can rightfully

### *Kant's Practical Deduction of Moral Obligation*

count to our theoretical knowledge, and that this is the premise underlying the whole argument. It cannot be denied that Kant makes statements that fit well with such an interpretation, both in the *Groundwork* and in the *Critique of Pure Reason*.<sup>14</sup> But irrespective of how these claims generally should be construed, it is at least plausible to read the present passage in the practical perspective. For at G 458 Kant continues:

By *thinking* itself into the intelligible world practical reason does not overstep its limits in the least: it would do so only if it sought to *intuit or feel itself* into that world.

The independence of reason that leads us into the intelligible world is, at least here, conceived as an operation performed in thinking, and it is practical, since thinking or judging is also acting.<sup>15</sup> In the remaining pages of the text of *Groundwork* III, Kant points out that we have no intuition of an intelligible order, and therefore no cognition thereof, but that it is an idea employed in acting. This suggests that the idea of an intelligible world should be compared to the practical idea of a kingdom of ends (without of course being exactly the same idea), rather than to speculative metaphysical cognition.

## 5. SOME FINAL CONSIDERATIONS

In reflecting on Kant's claims about the spontaneity of reason the distinction between the theoretical and the practical perspective becomes an important issue. The distinction is roughly that between cognition and action as different interests of reason. The architectonic interrelations between these perspectives, in particular with regard to the doctrine of the primacy of practical reason, is a crucial issue in

---

<sup>14</sup> See for instance *Critique of Pure Reason* A547/B575.

<sup>15</sup> The view that acting includes judging does not as such commit Kant to seek a theoretical proof for our spontaneity. See also Wood 1999, 178.

this connection. In this essay the notion of distinct perspectives has been used without further clarification, and I must leave it at that.<sup>16</sup>

It is now possible to give a rough summary of Kant's practical deduction of moral obligation. There is a theoretical part, showing that freedom is logically possible, which strictly speaking is not part of the deduction but rather a subsidiary argument, relying on the first *Critique*.<sup>17</sup> Its function is to open a space outside of the theoretical sphere. The principal part, the deduction of moral obligation, brings a content to this space. It shows how we as agents view ourselves under the idea of freedom, thereby regard ourselves as members of an intelligible world, and thus as autonomous (whereby the moral law follows); though we also perceive our actions as natural events and ourselves as part of nature. The basic task of the deduction, to make understandable why the moral law takes the form of an imperative for us, is solved by considering the consequences of our taking these two standpoints. The standpoint in which we consider ourselves as rational beings, members of an intelligible world, is recognized as the authentical one, because the other standpoint, in which we view ourselves as effects of the causality of nature, leaves no room for the self-determination taken for granted in acting. Therefore, in a practical perspective, the view of oneself as member of the intelligible world (as *Intelligenz*) is accorded the status of defining one's "proper self" (*eigentliche Selbst*) (G 457–458). What one would do as a purely rational being is now understood as one's own will as *Intelligenz*, and this will takes the form of an imperative for a being that is only partly rational. The obligation is self-imposed, because we recognize the rational will as our proper self.<sup>18</sup>

---

<sup>16</sup> The two Introductions to the *Critique of the Power of Judgment* should be good points of departure for investigating Kant's perspectivism. See also Buchdahl 1992 and Palmquist 1993.

<sup>17</sup> The practical deduction ends already at G 454, which is indicated by Kant when he proceeds to give an example that "confirms the rightness of this deduction," evidently meaning that the deduction is now completed. This is argued for in Schönecker 1999, 376.

<sup>18</sup> Schönecker (1999, 371) offers a metaphysical interpretation of the so-called onto-ethical principle (*ontoethischen Grundsatz*), embodied in Kant's dictum that "the intelligible world contains the ground of the sensible world and therefore also of its laws" (G 454). On the present view, this statement articulates the idea of the intelligible world in terms of the superiority of

### *Kant's Practical Deduction of Moral Obligation*

Finally a word on Kant's apparent change of mind in the *Critique of Practical Reason*. This is a complex issue that can only be addressed most tentatively here with a single point. It can hardly be denied that there are important differences between Kant's two accounts.<sup>19</sup> The main one is that instead of describing an entry to the realm of morality as in G III, Kant in the *Second Critique* assumes that we are already there – moral obligation is a *fact of reason*. Consequently, the very idea of a deduction of moral obligation is now superfluous. But since, according to G III, even “the most hardened scoundrel” (G 454–455) transfers himself into the higher realm, constrained to do so by the idea of freedom, moral obligation must present itself to him as a fact. The following summary of the result of the practical deduction of moral obligation can thus also be read as an early formulation of the doctrine of the fact of reason:

from this standpoint he is conscious of possessing a good will which, on his own admission, constitutes the law for the bad will belonging to him as a member of the sensible world – a law of whose authority he is aware even in transgressing it. The moral “I ought” is thus an “I will” for man as a member of the intelligible world; and it is conceived by him as an “I ought” only in so far as he considers himself at the same time to be a member of the sensible world (G 455).

---

reason over nature in a practical perspective. The “superiority in ontological status of the intelligible world” (Schönecker 1999, 379) should not be taken in a theoretical sense, which would only result in a transcendent metaphysical theory difficult to accommodate with the critical philosophy.

<sup>19</sup> For an interesting attempt to show their compatibility, see Högemann 1980, 291–293.

## *Acquisitio Originaria* and Epigenesis: Metaphors for the *A Priori*

It is safe to say that *a priori* conditions for cognition play a most important role in Kant's transcendental philosophy. Indeed, in the *Critique of Pure Reason*,<sup>1</sup> such cognition is called "transcendental" which pertains to "our mode of cognition of objects insofar as this is to be possible *a priori*" (B25), rather than to objects directly. Showing that there are such transcendental conditions for the possibility of experience, as well as identifying them as the forms of intuition (space and time) and the pure concepts of the understanding (categories), are crucial tasks of the critical undertaking. Kant makes it clear that his notion of the *a priori* is not to be confused with the rationalistic conception of innate ideas. He is as strongly opposed to the empiricist tradition which rejected innate ideas. In his response in 1790 to the criticism of J. A. Eberhard, Kant employed a legal analogy, labelling his distinctive view *acquisitio originaria*, an original (not empirically grounded) acquisition of the forms of intuition and the categories on the occasion of experience.<sup>2</sup> This is to be taken as a position independent of innatism as well as empiricism. The question poses itself as to what original acquisition is supposed to mean in a

---

<sup>1</sup> Henceforth referred to as CPR. Citations indicate the pagination of the first edition (A) and/or the second edition (B). For other works of Kant's, references are given to volume and page of the *Akademie* edition (Ak.).

<sup>2</sup> Ak. 8:221–223.

transcendental or epistemological<sup>3</sup> context, and how it differs from innatism.<sup>4</sup> This is the topic of section 1 of the present essay.

Section 2 discusses a biological analogy used by Kant when he explains the status of the categories. Concepts from the eighteenth century theory of generation (which more or less corresponds to today's embryology and developmental biology) are deployed in order to illustrate what options there are in accounting for the role of the categories. In CPR B166–168 three types of theory for explaining the origin and development of organization in a living being (*generatio aequivoca*, preformation, and epigenesis) are compared to three epistemological positions regarding the origin and status of the categories (empiricism, rationalism, and transcendental idealism). Kant aligns his own position to the epigenetic theory of generation. I shall take a closer look at these biological theories with the aim of clarifying the epistemological analogy. Since there also appears to be a close connection between the themes of epigenesis and *acquisitio originaria*, their relation will be examined for the sake of elucidating the status of the *a priori* in Kant's philosophy.

In addition to its use as analogy, Kant was interested in the theory of generation as a part of his exploration of philosophical issues in biological science. In the *Critique of the Power of Judgment* (CJ) he classified the various theories on offer, and here too he endorsed epigenesis, but now in a purely biological context (CJ 422–424). An examination of Kant's assessment of this biological debate might contribute to the understanding of his epistemological views, by making the terms of the analogy more accessible. But it also opens some intriguing questions about the nature of the parallel drawn between philosophy and biology. Does Kant use biological terms purely metaphorically or analogically, with the sole purpose of illustrating his epistemological tenets, or is there a deeper connection between epigenesis with regard to the categories and epigenesis as a theory of generation?

Section 3 examines whether it is possible to connect Kant's references to epigenesis in epistemology and biology to a unified

---

<sup>3</sup> The terms “epistemological” and “epistemic” will be used as equivalent to “transcendental”.

<sup>4</sup> Important studies of the concept of *acquisitio originaria* against the background of the debate on innatism are Petzäll 1933 and Oberhausen 1997.

biological theory of cognition. In a recent essay Phillip Sloan boldly proposes such a biological reading of Kant's epistemology. He shows that biological terminology is frequently present in Kant's explanations of his philosophical position, and he argues that this has to be understood from the point of view of innatism, so that Kant's epistemology turns out to be a theory of biological inheritance of *a priori* concepts.<sup>5</sup>

Though some attempts have been made before to transform Kant's philosophy along the lines of biological innatism (for instance by Konrad Lorenz), what Sloan proposes is an *interpretation* of Kant in such terms (and not just a position loosely inspired by Kant). As such, it is hard to see how it is to be reconciled with Kant's stress on the necessity of the transcendental conditions. In addition to this fundamental problem for attempts to unite Kant's strong commitment to necessity with biological innatism, some of Sloan's readings of particular passages of Kant in philosophical as well as biological contexts are rather implausible, as I shall try to show.

The concluding remarks note another biological analogy frequently employed by Kant, that between the systematicity of reason and the functional integration of the parts in an organism.<sup>6</sup> Kant's teleological conception of reason is often described in terms of biological organization. While this could be thought to strengthen the case for an interpretation along the lines of biological innatism, I shall argue that a consideration of Kant's views on teleology in biology actually points in the opposite direction. The teleological notion of organism appealed to by Kant is not straightforwardly empirical. Though occasioned by experience, and certainly applied to empirical objects in biological science, it expresses a way of judging which makes use of a normative point of view. While this teleology is acceptable and even indispensable in biology, it cannot be accorded a constitutive status from a philosophical perspective. Biology is thus a rather special science according to Kant, as it is constituted by means of a regulative maxim (teleology) that pertains to the reflecting power of judgment, which has more to do with how we have to interpret nature given our cognitive peculiarities than with what we can objectively ascribe to nature (see CJ 185–186). Sloan's idea of a

---

<sup>5</sup> Sloan 2002.

<sup>6</sup> See for instance CPR Bxxxvii–xxxviii.

grounding of the categories in natural science is problematic already in view of Kant's stress on the opposite direction of the grounding relation, but considering the merely regulative status of biological teleology, the prospects for grounding the categories in the biological part of natural science are even worse.

On Kant's view, the teleological structure of reason cannot be derived naturalistically from biological facts; rather, the possibility of biology as a science relies on an analogical use of the concept of purposive production, which pertains to reason. Hence one has to conceive of *a priori* concepts as having their source in reason (in a broad sense which includes the understanding), and reason must be presupposed as *sui generis*. This is the sense of Kant's description of his position as "a system of the epigenesis of pure reason" (B167). Difficult metacritical questions can be posed concerning what grounds there are for the very assumption of something such as reason. I shall not attempt to go deeply into this issue, but only point to Kant's view that any argument ultimately appeals to reason, so that even a skepticism that denies reason actually tries to "prove by reason that there is no reason."<sup>7</sup> Similarly, for Kant to try to reduce *a priori* principles such as the categories to anything passively given, like innate ideas, would be to put the cart before the horse. Kant's conception of reason as a teleological structure cannot be construed as referring to a theoretical entity posited on the basis of empirical findings in biology; reason is rather a presupposition for achieving objectivity in philosophy and science.

## 1. INNATISM AND *ACQUISITIO ORIGINARIA*

Kant frequently criticizes empiricist accounts of cognition. His basic objection is that empiricism can never achieve more than a mere "rhapsody of perceptions," as it neglects the *a priori* forms of experience (CPR A156/B195–196). With regard to sensible intuition, something is needed "which allows the manifold of appearance to be ordered in certain relations" (B34). Sensations are received in space and time, and these latter are not empirically derived; they are rather

---

<sup>7</sup> *Critique of Practical Reason*, Ak. 5:12.



forms of intuition, *a priori* conditions of receptivity (A26/B42). For the possibility of judgment, a further set of forms are required, namely the categories or pure concepts of the understanding. These are likewise *a priori* (not derived from experience), and they bring forth cognition, expressible in judgments, of what is given in space and time (compare A77–78/B102–104).

That sensations alone are insufficient for cognition is an old idea associated with the rationalistic tradition. It was standardly connected with innatism, according to which the mind possesses innate concepts that are predisposed to match experience. It might seem natural to link Kant to the rationalist tradition and consider his *a priori* forms to be such innate ideas, or something very close to that.<sup>8</sup> Two main varieties of innatism can usefully be distinguished. The first one (associated with Descartes) assumes that there are inborn representations that lie ready in the mind and are open to conscious inspection, whereas the second (Leibnizian) version denies this and merely asserts that it is the mind's dispositions or ways of operating that are innate.<sup>9</sup> It is the latter type of innatism that has most frequently been associated with Kant's view, often in conjunction with the hypothesis that Leibniz' *Nouveaux essais*, published some years before the critical turn in Kant's thinking, is the source from which he got the idea.<sup>10</sup>

There are passages in the *Transcendental Aesthetic* of CPR suggestive of innatism with respect to the forms of intuition. Consider

---

<sup>8</sup> See for instance Nagel 1983, 49: "Kant does not describe our spatial ideas as *innate*, but his description of them as *a priori* amounts practically to the same thing."

<sup>9</sup> In the following I will label these versions of innatism the Cartesian and the Leibnizian. The actual views of Descartes and Leibniz are of course more complex. Petzäll (1933, 10–14) traces a development in Descartes from viewing innate ideas as conscious contents given before experience to viewing them as dispositions. Perhaps a dispositional view was what Descartes had in mind all along. And in Leibniz there is a variety of formulations of the doctrine that, according to Petzäll (1933, 18–24), can be unified only in the context of the ontology of preestablished harmony. See also Nicholas Jolley (1988, 86–87) who, following C. D. Broad, argues that Leibniz' dispositions can be understood only as unconscious perceptions present in the mind.

<sup>10</sup> See for instance Kemp Smith 2003 [1923], 92.

the following examples: the form of appearance must “lie ready [for its matter] in the mind *a priori*” (A20/B34); space “must be encountered in us *a priori*, i.e., prior to all perception of an object” (B41); “the form of all appearances [is] given in the mind prior to all actual perceptions, thus *a priori*” (A26/B42).

The pure concepts of the understanding are also described in ways reminiscent of the rationalistic view on innate ideas. In A66/B91, Kant declares that the proper task of transcendental philosophy is to “pursue the pure concepts into their first seeds and predispositions in the human understanding, where they lie ready.”

The pure concepts are likened to seeds or germs (*Keime*) and predispositions (*Anlagen*), and the passages from the *Transcendental Aesthetic* are likewise easy to interpret as expressing innatism, but in addition they seem to point to its Cartesian version, rather than to the Leibnizian one. Space and time are referred to as if lying ready in the mind waiting for perceptions, and the *a priori* merely appears to denote temporal priority.

Cartesian innatism concerning space and time is not a very attractive position; perhaps it is not even intelligible. What does it mean to say that space and time in their infinity are present in the mind even without any empirical content, and why would such a view be any better than the Newtonian realism about which Kant says that it assumes “two eternal and infinite self-subsisting non-entities [*Undinge*]” (A39/B56)? And how could it make sense to speak of time (as form of intuition) being given *before* perceptions? If time is something which has its existence in the mind and is there even in the absence of perception, how could it be *temporally* related to something outside of time (the advent of perceptions), considering that temporal relations are in time?

There are good reasons to reject this picture and interpret the priority of forms of intuition in an epistemological rather than temporal sense.<sup>11</sup> At the very beginning of the *Introduction* to the second edition of CPR, Kant says that “[a]s far as time is concerned [...] no cognition in us precedes experience” (B1). If time lay ready in the mind, we would thus have no cognition of it without experience (which involves perception); this seems difficult to reconcile with

---

<sup>11</sup> The following passages, as well as others pointing in the same direction, are cited in Falkenstein 1995, 84–85. See also Paton 1936, 77–80 and 138.

Cartesian innatism. The passage continues with a discussion of the *a priori* as prior to experience in a nontemporal sense,<sup>12</sup> which shows that apriority is not a matter of being present before the *a posteriori*. The inappropriate view (engendered by Kant's own way of expressing himself) of experience as the result of combining pre-given space and time with separately given perceptions is repudiated in A429n/B457n: "empirical intuition is not put together out of appearances and space (out of perception and empty intuition). The one is not to the other a correlate of its synthesis, but rather it is only bound up with it in one and the same empirical intuition, as matter and its form."

If this speaks against Cartesian innatism with regard to space and time, does it not also rule out dispositional (Leibnizian) innatism? Before drawing that conclusion, we must take into account other works of Kant where he takes a stand on the question of innatism. The relevant works are the *Inaugural Dissertation*<sup>13</sup> (1770), standing half-way between the pre-critical and the critical Kant, and the response to Eberhard<sup>14</sup> (1790). Common to these works are statements to the effect that forms of intuition as well as pure concepts are neither innate nor acquired in the ordinary sense, but originate from a special kind of acquisition due to the action of the mind, and that what is innate is not *a priori* representations, but their ground in the cognitive capacities.

In the *Inaugural Dissertation*, the account of space and time is in many respects similar to that in the *Transcendental Aesthetic* of CPR. It is argued that space and time are formal principles of the sensible world, that is, for things considered as phenomena, and that they are pure intuitions. Kant then raises the question whether they are innate or acquired. "The former view [...] ought not to be that rashly admitted, for it paves the way for a philosophy of the lazy [...] which, by appealing to a first cause, declares any further enquiry futile."<sup>15</sup>

---

<sup>12</sup> B1-B2.

<sup>13</sup> *On the Form and Principles of the Sensible and the Intelligible World*.

<sup>14</sup> Its title is *On a Discovery Whereby any New Critique of Pure Reason Is to Be Made Superfluous by an Older One*. Eberhard had attacked Kant on the ground that the putative novelties of the CPR, when not false, were already known in the Leibniz-Wolff tradition.

<sup>15</sup> Ak. 2:406.

But though acquisition is the only remaining alternative, this cannot be a case of ordinary empirical acquisition, since Kant has argued for the universality of formal principles. Principles of form contain “the ground of the universal connection” of the matter, in this case sensation; matter itself cannot provide such universality.<sup>16</sup> Space and time have rather been acquired “from the very action of the mind, which coordinates what is sensed by it, doing so in accordance with permanent laws. [...] Nor is there anything innate here except the law of the mind, according to which it joins together in a fixed manner the sense-impressions made by the presence of an object.”<sup>17</sup> What is said to be innate is thus not the representations of space and time (which would be Cartesian innatism), but a law of the mind for coordinating sensations (which at least looks like Leibnizian innatism).

Kant’s view in the *Inaugural Dissertation* on pure intellectual concepts is much less developed than the later doctrine of the categories, and remains in many respects within the rationalistic tradition.<sup>18</sup> It allows the possibility of metaphysical cognition of things as they are by means of the “real use of the intellect,”<sup>19</sup> but it also (like CPR) denies human beings intellectual intuition,<sup>20</sup> which makes it entirely unclear how the real use of the intellect is supposed to provide cognition.<sup>21</sup> At any rate, Kant lists some concepts belonging to the nature of the intellect,<sup>22</sup> and indicates that their

---

<sup>16</sup> Ak. 2:398. This is closely connected to the thesis in CPR that what is *a priori* is characterized by a necessity and universality which cannot be reached by a comparative generality grounded in experience (B3).

<sup>17</sup> Ak. 2:406.

<sup>18</sup> In the *Inaugural Dissertation*, written in Latin, Kant did not distinguish between understanding and reason, but used only the term *intellectus*. I follow Falkenstein’s translation of it as “intellect” (though not his rendering of *Verstand* as “intellect” in CPR). See Falkenstein 1995, 367n 3.

<sup>19</sup> Compare Ak. 2:393–394.

<sup>20</sup> Ak. 2:396.

<sup>21</sup> Kant also says (Ak. 2:395–396) that the intellect has a “dogmatic end” that issues in a paradigm of “noumenal perfection.” This notion, reminiscent of the Ideal of pure reason in CPR, points forward to the later view on the regulative function of reason.

<sup>22</sup> Ak. 2:395. The concepts mentioned are possibility, existence, necessity, substance and cause; all of which are among the categories of CPR.

origin is neither empirical nor innate, but that they are “abstracted from the laws inherent in the mind.”<sup>23</sup>

So both with regard to sensible and intellectual cognition, in his *Inaugural Dissertation* Kant refers to laws “inherent in the mind”, rather than innate ideas, as the proper place of origin for *a priori* representations. These laws themselves, as they are the mind’s way of functioning, can be seen as innate dispositions for bringing about *a priori* representations, which amounts to Leibnizian dispositional innatism.

This picture is even clearer in the response to Eberhard of 1790. Kant here claims that his critical philosophy admits no innate representations, but considers all representations to be acquired. He adds, however, that “there is also an original acquisition” of space and time and of the unity brought about by the understanding. “There must indeed be a ground [...] in the subject [...] which makes it possible that these representations can arise in this and no other manner, and be related to objects which are not yet given, and this ground at least is *innate*.”<sup>24</sup> What is innate is thus not the representation of space, but its ground: mere receptivity as such. Likewise the pure concepts of the understanding are acquired, “though their *acquisitio*, like that of space, is no less *originaria* and presupposes nothing innate save the subjective conditions of the spontaneity of thought.”<sup>25</sup>

What Kant suggests is that the *a priori* conditions of experience (sensible and intellectual) are not innate ideas, but acquired from receptivity and spontaneity, the dual basis on which the transcendental theory of cognition rests. These faculties themselves are innate dispositions to produce the *a priori* representations. Apparently this conforms to Leibnizian innatism. The term used for indicating that *a priori* representations have this origin is “*acquisitio originaria*.” It is a legal term, pertaining to property right. When something is owned by someone and then changes owner, the acquisition is called derivative. This contrasts to original acquisition, which takes place when something is acquired which before the

---

<sup>23</sup> Ak. 2:395.

<sup>24</sup> *On a Discovery*, Ak. 8:221.

<sup>25</sup> Ak. 8:223.

acquisition did not belong to anyone.<sup>26</sup> The point of the analogy with these legal notions is to stress that neither space and time nor the pure concepts are derived from the objects as they are in themselves, “rather [our cognitive faculty] brings them about, *a priori*, out of itself.”<sup>27</sup>

As there are obvious similarities here between the precritical *Inaugural Dissertation* and the late response to Eberhard, and the former work in many ways points forward to the *Critique of Pure Reason* whereas the latter is meant as a defence of the critical doctrines, it is natural to conclude that *acquisitio originaria* is an unchanging feature of Kant’s philosophy from 1770 onwards. Furthermore, the admission of innate laws of the mind is often taken to show that Kant’s critical doctrine is based on a Leibnizian dispositional innatism.<sup>28</sup>

There are some reasons to doubt these seemingly obvious conclusions, however. To begin with the *Inaugural Dissertation*, it is not only in its account of the intellect that it differs from the later critical doctrine, but also, as is sometimes remarked, with regard to sensibility.<sup>29</sup> Only here does Kant speak of space and time as issuing from the mind’s “co-ordinating” or “joining together” sensations.<sup>30</sup> In CPR such talk of active operations is replaced by the notion of a mere receptivity, which receives objects in a spatial and temporal manner, whereas synthesis, as activity, is ascribed to the understanding (though effectuated by the imagination), and presupposes a manifold received in sensibility. Though this difference between the *Dissertation* and Kant’s later view is not immediately relevant to the issue at hand (as it regards only the question whether sensibility is a passive capacity), it serves as a reminder that it cannot be taken for granted that even those views in the *Dissertation* and the first *Critique* that closely resemble each other are exactly the same.

---

<sup>26</sup> On the history of this distinction, see Oberhausen 1997, 121–128.

<sup>27</sup> Ak. 8:221.

<sup>28</sup> It should be noted that these conclusions are not identical: *acquisitio originaria* could be a permanent doctrine in Kant’s critical philosophy but yet not be a kind of Leibnizian innatism. This is Oberhausen’s view (1997).

<sup>29</sup> See for instance Kemp Smith 2003 [1923], 92, Klemme 1996, 57–59, and Falkenstein 1995, 93 and 96.

<sup>30</sup> Ak. 2:403, 406.

On the positive side, the *Dissertation* clearly was an attempt to articulate a new view on issues such as *a priori* cognition and innatism, and this orientation ultimately led to the critical doctrine of transcendental conditions for the possibility of experience. Though there are similarities to Leibniz' view in the *Nouveaux essais*, published a few years before the *Dissertation*, and an influence is plausible to assume, it cannot simply be taken as a fact that Kant was strongly influenced by this work.<sup>31</sup> Regardless of whether there actually was such an influence, there are reasons to consider the resulting view to be something quite different from Leibnizian dispositional innatism. For even though Kant may have borrowed the terms in which to describe his position (and perhaps even got the idea for it) from the *Nouveaux essais*, the view he was moving toward in the *Dissertation* and developed more fully in CPR is epistemically oriented, and appeals to the notion of transcendental conditions for experience, whereas the Leibnizian view is closely bound to the ontology of monadology.<sup>32</sup> This makes it problematic to ascribe dispositional innatism to Kant, even though he couches his theory in terms pertaining to that doctrine.

When it comes to assessing Kant's response to Eberhard in a larger context, some caution is also called for. As Petzäll (1933, 39–40 and 46–52) and Oberhausen (1997, 92–93) have stressed, Kant's use of innatist locutions is conditioned by the terms set for the discussion by his adversary. It is Eberhard who demands that Kant must position his critical philosophy with respect to the traditional alternatives innate or acquired. The debate between Eberhard and Kant was conducted in a very rude manner, and apparently Kant wanted to show that he understood Leibniz better than the Leibnizian Eberhard did. In order to accomplish this goal, he tones down the differences between his critical philosophy and Leibniz' views (for instance concerning the status of sensibility), and offers an interpretation of Leibniz according to which ontological assumptions

---

<sup>31</sup> Oberhausen (1997, 120) points out that this influence has never been proven, despite what many commentators have claimed. (Though this lack of proof doesn't prove much either, of course.)

<sup>32</sup> I cannot here provide an examination of Leibniz thought to back up these claims. But see Petzäll (1933, 18–24), who argues that Leibniz' conception of innate dispositions is dependent upon his view of the soul as a substance.

about the substance grounding the innate dispositions should not be taken literally.<sup>33</sup> In view of this, it seems somewhat over-hasty to read Kant's response to Eberhard's question on innatism as a final statement on the nature of apriority in the critical philosophy.<sup>34</sup>

To return now to CPR, its main thrust is epistemological rather than ontological. The fundamental standpoint of transcendental idealism with its distinction between appearance and things in themselves leads to a conception of the *a priori* as necessary conditions for cognition, in contrast both to a psychologically given content (Cartesian innatism) and to dispositions for obtaining cognition requiring an ontological harmony pre-given in the substantial soul (Leibnizian innatism). In § 36 of *Prolegomena*, C. A. Crusius' particular version of innatism, which (as interpreted by Kant) presupposes a harmony between cognition implanted in the mind and extra-mental nature, is criticized on the ground that it lacks a criterion for deciding whether an idea that we possess is veridical or spurious.<sup>35</sup> Not only is the theological assumption of rules implanted by a deity gratuitous, but furthermore, even if it were accepted it would be of little use, since any idea could be taken to be such an implanted one in the absence of distinguishing criteria providing some reason to take the idea as veridical. Suppose some such criteria are proposed, for instance necessity and universality (which characterize the *a priori* according to Kant). Then we have moved to an epistemic level of investigation, where we search for necessary and universal conditions for cognition, and considerations of the mind's content or the dispositions inherent in the soul become irrelevant.

It is admittedly not an easy task clearly to separate the transcendental from the psychological and ontological levels of inquiry, as is shown by the perennial difficulties for Kant's

---

<sup>33</sup> Ak. 8:249, compare Petzäll 1933, 50-51.

<sup>34</sup> In a study on the Kant-Eberhard controversy, Manfred Gawlina also concludes that Kant's reading of Leibniz in *On a Discovery* is tactically motivated (Gawlina 1996, 292). According to him, the passages where receptivity and spontaneity are referred to as innate amount to a naturalization of epistemological concepts that has no counterpart in CPR, and this deviation must likewise be explained as being due to tactical considerations (Gawlina 1996, 261).

<sup>35</sup> Ak. 4:319.



commentators of justifying the assumptions made in transcendental philosophy about the cognitive faculties, and to explain the relation between the transcendental subject and individual persons. Though I will not here enter this large issue, it seems obvious that Kant, whether successfully or not, aimed to establish the validity of a transcendental level of inquiry, and this very attempt is enough to distinguish the critical philosophy from Cartesian and Leibnizian innatism.

The point can also be put like this: Kant's critique starts from the separation of two fundamental sources of cognition, receptivity and spontaneity. The question as to how these capacities are to be related to an individual person's cognitive psychology is hard enough. To call receptivity and spontaneity "innate," as Kant is moved to do in the context set by Eberhard's questioning, does not add anything to clarifying their status above the description of them as fundamental capacities, but threatens to bring irrelevant concerns from earlier philosophical traditions into the discussion.<sup>36</sup>

## 2. THE EPIGENESIS OF PURE REASON

Kant's *Critique of Pure Reason* is replete with metaphors and analogies. Many are taken from the legal domain. For instance, we find the critical undertaking compared to a court of justice (*Gerichtshof*) instituted by reason,<sup>37</sup> which goes well with the aim of determining the conditions for possible experience and exposing illegitimate claims of metaphysical knowledge. Kantian key concepts such as *deduction*, *antinomy*, *regulative* and *constitutive* also have a legal origin.<sup>38</sup>

---

<sup>36</sup> Compare Yamane 2001, who stresses the radical difference between Kant's project and the traditional debate on innate ideas, pointing out that *originaria* corresponds to *ursprünglich* which is a key term in Kant's transcendental philosophy.

<sup>37</sup> CPR Axi–xii.

<sup>38</sup> On *antinomy*, see Hinske 1965, 488–489. On *regulative* and *constitutive*, see Tarbet 1968, 269. That *deduction* (as the term is used in the transcendental deduction of the categories) carries the legal sense of justification

Other metaphors used with some frequency in CPR relate to battle, geography, biology, geometry and architecture.<sup>39</sup> I shall here examine the biological analogy made in § 27 of the B-edition of the transcendental deduction, where Kant describes possible ways to construe the agreement between categories and experience, and compares these options to the alternatives in the theory of generation (that is, the development of an organism from germ to adult state). The relations between the biological theories bear a close structural resemblance to the relations between the legal concepts deployed in the *acquisitio originaria* discussion, and as the change from a legal to a biological analogy is a change to a domain that could well be most relevant to at least a psychological version of innatism, it is interesting to take a closer look at what Kant has to say here. In addition to that, Kant examined questions on the foundations of biological theory, and his writings on this subject (especially in the third *Critique*) contain discussions on theories of generation that could broaden the investigation.

In the B version of the transcendental deduction of the categories, § 27 is the final section. Its role is not to furnish new premisses for the argument of the deduction, but to summarize what has already been shown and to present the conclusions with the aid of an analogy. The analogy used is that to the theory of generation, and the presentation of it takes up most of § 27. The context in which it is introduced is the question of how concepts can be thought to have a necessary agreement with experience:

---

of a claimed right is mentioned at A84/B116. This was noted by Heidegger 1991 [1929], 85–86, Paton 1936, 313, and Taret 1968, 266–267; yet it seems to have become generally acknowledged only after the account of the term’s history given in Henrich 1989. But see also Kuehn 1997, 245, who argues that Henrich takes the parallel with legal deductions too literally, thereby downplaying the argumentative rigour of Kant’s deduction. A striking example of how Kant fuses the legal and the logical senses of “deduction” is found in his 1785 essay *On the Wrongfulness of Unauthorized Publication of Books*, where a (clearly legal) “deduction of the right of a publisher against an unauthorized publisher” is given as a syllogistic proof (Ak. 8:79–80; Kant 1996, 29).

<sup>39</sup> Some of these metaphors are discussed in Taret 1968. For the geometrical ones see Palmquist 1993, 17–21.

either the experience makes these concepts possible or these concepts make the experience possible. The first is not the case with the categories (nor with pure sensible intuition); for they are *a priori* concepts, hence independent of experience (the assertion of an empirical origin would be a sort of *generatio aequivoca*). Consequently only the second way remains (as it were a system of the *epigenesis* of pure reason): namely that the categories contain the grounds of the possibility of all experience in general from the side of the understanding. [...] If someone still wanted to propose a middle way between the only two, already named ways, namely, that the categories were neither *self-thought a priori* first principles of our cognition nor drawn from experience, but were rather subjective predispositions for thinking, implanted in us along with our existence by our author in such a way that their use would agree exactly with the laws of nature along which experience runs (a kind of *preformation-system* of pure reason), then (besides the fact that on such a hypothesis no end can be seen to how far one might drive the presupposition of predetermined predispositions for future judgments) this would be decisive against the supposed middle way: that in such a case the categories would lack the *necessity* that is essential to their concept. (B166–168)

Before going into the biological connotations of the terms, I shall briefly delineate the epistemological issues.<sup>40</sup> To begin with the preformation-system (which for Kant is no genuine but only a “supposed” option), it is obvious that it corresponds to Crusius’ view, criticized in *Prolegomena* § 36. A general flaw in this sort of view is that its appeal to a deity that has implanted the dispositions in the subject opens the way for irresponsibly multiplying the number of judgments predisposed for. But the decisive argument against preformation is that subjective dispositions have no necessity and

---

<sup>40</sup> The following account largely agrees with the ones in Zöllner 1988, Ingensiep 1994, and Haffner 1997, 160–172. But whereas Zöllner stresses that epigenesis is the generation of the system of categories as *a priori* philosophical cognition, Haffner (1997, 168) points to the dual sense of the epigenesis of reason: it regards both the generation of the categories and the constitution of experience as a product.

cannot therefore correspond to the *a priori* character of the categories. For genuine necessity Kant requires a tighter connection between representation and object than mere correlation, even if divinely ordained. One might argue that subjective dispositions possess such necessity, given that their harmony with experience is preordained by God. Presumably Kant's answer would be that God's existence being unprovable (something CPR argues for at length), the putative harmony is a gratuitous hypothesis and consequently no necessity can be claimed.

The serious alternatives for explaining how concepts can agree with experience necessarily and not just contingently are that experience makes the concepts possible or that the concepts make experience possible. The first one is the empiricist explanation, according to which experience grounds the possibility of concepts. This is plausible for empirical concepts, but when it comes to categories (and also pure intuitions), their status as *a priori* excludes any empirical origin; experience can never give genuine necessity, but only a comparative (inductive) generality devoid of necessity (B3–4). One might now wonder why this alternative, as it is incapable of explaining the relation between *a priori* concepts and experience, should be considered a more serious option than Crusius' theory of preformation. After all, according to Kant both theories share the fault of not being able to account for the necessity pertaining to the categories. So why is there but two possible ways, whereas the middle way between them is merely "supposed"? It appears that Kant considers the view according primacy to experience respectable because what it states can be investigated. In many cases (in regard to empirical concepts) it actually is the case that experience makes the concepts possible; but it so happens that categories are different in view of their apriority and therefore cannot be explained in this way. Hence, empiricism is not always wrong. The preformationist view, on the other hand, is *never* a serious option since it appeals to intellectual laziness in bringing in a *deus ex machina*,<sup>41</sup> and there is no instance where it constitutes a genuine explanation. To be sure, in the special case of the categories, an empirical ground is as impossible an explanation as preformation is. It is then compared to *generatio aequivoca*, a theory in biology which is, as we will see, highly

---

<sup>41</sup> This theme in Kant is surveyed in Petzäll 1933.

unsatisfactory for Kant. But in other cases, explaining concepts out of experience makes sense. A further difference is that preformationism, with its subjective dispositions, can only speak of what a subject must think, not about any objectively valid judgments (that is, valid with respect to objects), and therefore invites skepticism (B168). In contradistinction, the possibility to base empirical concepts on experience depends on empirical laws (for instance those pertaining to causal connections in perception), although these laws, to obtain the necessity implied by lawfulness, presuppose categories that are independent of experience. That laws play a role here might explain why also the empiricist option is considered as one way “in which a *necessary* agreement of experience with the concepts of its objects can be thought” (B166), even though empiricism cannot explain the necessity in the categories. Preformationism locates the laws on the supernatural plane, whereas empiricism and transcendental philosophy attempt to explain lawfulness within the empirical world.

The remaining possible explanation for the agreement between concepts and experience is that experience is made possible by the concepts. This is the case for the categories, for they contain the grounds provided from the understanding for the possibility of experience. The categories are “self-thought” principles of our cognition (B168): this Kant calls the epigenesis of pure reason. There is a clear parallel between epigenesis and *acquisitio originaria*, in that both of these notions indicate that representations are produced from the side of the cognitive capacity, and neither from experience nor from innately preformed representations.

The parallel is explicitly drawn in some of Kant’s *Reflexionen*. For instance, in R4851,<sup>42</sup> dated to the 1770s, a distinction is made between a concept considered as “educt” or as “product,” and the former is linked to preformation, the second to epigenesis. In the

---

<sup>42</sup> Ak. 18:8. “Ob die Begriffe blos *educta* oder *producta* sind. praeformation und *epigenesis* (*producta* entweder durch physischen (empirischen) Einfluss oder durch Bewusstseyn der formalen Beschaffenheit unserer Sinnlichkeit und Verstandes bey gelegenheit der Erfahrung, mithin doch *producta a priori*, nicht *a posteriori*.) Die Lehre von *ideis connatis* führt zur Schwärmerey, *acqvisitae* sind *a priori* oder *a posteriori acqvisitae*, jene sind nicht immer intellectuell. Also ist die Eintheilung der Erkenntnis in sensitive und intellectuelle nicht die erste, sondern in die *a priori* oder *a posteriori*. Die erstere entweder sinnlich oder intellectuell.”

same remark a distinction between *a priori* and *a posteriori* acquisition is made, and *a priori* product and *a priori* acquisition are linked to each other. The doctrine of innate ideas is put in correspondence to preformation, and is dismissed as *Schwärmerei*. A further thing to note in this compressed note is that two possibilities are distinguished with regard to product (and thereby epigenesis): through empirical influence (*a posteriori*), or through “consciousness of the formal character of our sensibility and understanding on the occasion of experience” (*a priori*). It thus appears that what is called *generatio aequivoca* in CPR is here considered to be one of the types of epigenesis, namely the *a posteriori* one. To get a clearer view of the relations between these alternatives, and to understand terms like “epigenesis,” “educt,” and “*generatio aequivoca*,” we need to turn to Kant’s writings on the theory of generation in the *Critique of the Power of Judgment*.

In CJ § 80, Kant claims that even though the similarities between different species suggest that they have a real kinship (a common descent), the origin of life cannot be understood as an entirely mechanical process on the basis of “raw matter” alone (CJ 418–419). A genealogical hypothesis that connects all organisms in a great family Kant labels the “archaeology of nature” (CJ 419, 428). He considers it to be a somewhat daring hypothesis, an “adventure of reason,” but in no way absurd. The problem with the hypothesis is just that it has no empirical support (CJ 419).<sup>43</sup> This genealogical hypothesis is absurd only if combined with the assumption of *generatio aequivoca*, “the generation of an organized being through the mechanism of crude, unorganized matter”,<sup>44</sup> as long as *generatio*

---

<sup>43</sup> It might seem strange that Kant in this same passage adduces an empirical example: “as when certain aquatic animals are gradually transformed into amphibians and these, after some generations, into land animals.” However, the example is not intended as a statement of historical fact, but rather as something that could conceivably happen. The German “*wenn*” corresponds to “if” rather than to “when” – contrary to the translations of Guyer and Matthews (Kant 2000) as well as Pluhar (Kant 1987) – and “transformed” (“*ausbildeten*”) is, like the surrounding context, in the subjunctive mood, so that the rendering should be: “as if certain aquatic animals were gradually transformed.” Lovejoy (1911, 48) gets this right.

<sup>44</sup> According to McFarland (1970, 39), Kant accepted *generatio aequivoca* in the case of lower animals. But this misrepresents CJ 411–412, where Kant

*univoca* is accepted, according to which something organic is “generated out of something else that is also organic,” the hypothesis is admissible (CJ 419). In view of the special properties of organisms, which make teleological attribution inevitable, Kant takes reductive explanation to be inconceivable. He does not dogmatically rule out that organisms could have a mechanistic origination, but he sees such ultimate ontological cognition as beyond our epistemic reach.<sup>45</sup>

Mechanistic explanation of organisms is parallel to the empiricist account of *a priori* cognition. For Kant, such an account fails to explain the necessity and lawfulness of the *a priori*, in analogy to how *generatio aequivoca* fails to explain biological order.

The concept of *generatio univoca* is further divided into *homonyma* and *heteronyma*. The archaeology of nature would require *generatio heteronyma* (organisms giving birth to organisms of different species), but that is “nowhere to be found.” Experience shows examples of *generatio homonyma*, where the product “is in its organization itself homogeneous with that which has generated it” (CJ 420). This distinction is of some historical interest in the context of prefigurations of Lamarck’s and Darwin’s theories, but for present purposes it need not be pursued further.

In § 81 additional distinctions are introduced, beginning with that between *occasionalism* and *pretabilism*. These are positions that accept an irreducible teleological basis for generation, and therefore fall under *generatio univoca* (the generation of organic beings out of other organic beings). Unlike the *homonyma-heteronyma* distinction which concerned the offspring’s likeness to its parents, this division (named after the metaphysical doctrines of occasionalism and pretabilism) has to do with the question of when a non-mechanistic influence is supposed to have taken place.

According to occasionalism “the supreme cause of the world would [...] on the occasion of every copulation, directly give the

---

discusses the implications of assuming the spontaneous generation of a maggot, without himself making that assumption; compare Bommersheim 1927, 306. One passage in CPR (A835/B863) could be taken to show that Kant accepted *generatio aequivoca* in 1781. But most likely he employs this notion only for illustrative purposes.

<sup>45</sup> See CJ 400.

mingling matter its organic structure” (CJ 422).<sup>46</sup> Since such a doctrine denies natural beings any genuine reproductive capacity, making everything dependent on direct supernatural intervention, it is contrary to the very attempt of rationally assessing the phenomenon, and hence “no one who cares anything for philosophy will assume this system” (CJ 422).

According to prestabilism, it is only at the original creation of organisms that they have been given the predispositions for perpetuating the species. Prestabilism, in Kant’s further classification, takes either of two forms, individual or generic preformation (CJ 423). Individual preformation (often simply referred to in the history of science as the theory of preformation) is also called the theory of evolution (which was its standard name at the time). The word is here used according to the literal meaning of *evolutio*, the unfolding of something already existing, and thus has nothing to do with evolution in the modern sense.<sup>47</sup> The opposing theory of generic preformation is called the system of epigenesis.

The theory of individual preformation claims that either the sperm or the ovary contains a more or less complete organism, the parts of which have only to grow and unfold. The difficulty of explaining generation is solved, or rather avoided, by this postulated pre-existence. No organism is produced anew, each individual must have existed already at creation, encapsulated in another tiny pre-existing organism, in its turn enclosed within another one, and so on in a series of encasements.<sup>48</sup> Against this theory, Kant first points out that it relies as heavily on the supernatural as does the occasionalist theory from which it wants to separate itself: “as if it made a difference

---

<sup>46</sup> Pluhar’s translation (Kant 1987).

<sup>47</sup> Kant amasses yet more names for the theory of individual preformation in this same paragraph: “encapsulation” and “theory of involution”. Thus involution and evolution are not different theories, as assumed in an utterly confused editorial note in Kant 1992; see page 434n 60.

<sup>48</sup> Compare Leibniz’ enthusiastic reception of the theory of preformation: “Persons very accurate in experiments have already in our day perceived that it may be doubted whether an altogether new animal is ever produced, and whether animals wholly alive as well as plants are not already in miniature in germs before conception. [...] This again reveals to us hitherto unthought-of marvels of divine contrivance.” (*Considerations on the Principles of Life, and on Plastic Natures*, 1705; Leibniz 1951, 195.)



whether [beings were created] at the beginning of the world or during its course” (CJ 423).<sup>49</sup> Seen from the point of view of parsimony, occasionalism is even the preferable theory, since preformation requires the individual creation of a very great number of preformed beings that never develop, whereas occasionalism only assumes divine actions for those cases where organisms actually develop (CJ 423). Preformation also needs the assumption of a “multitude of supernatural arrangements that would be necessary in order to preserve uninjured the embryos formed at the beginning of the world and to save them from injury by the destructive forces of nature during the long time until their development” (CJ 423). On the empirical side, Kant notes that the existence of hybrids (offspring of parents from different species) is inconsistent with the theory of preformation (CJ 424). For if there is in the ovary a preformed being of a certain species, how can it take on traits from another species? The difficulty for the preformationists to explain the occurrence of hybrids reported by Kant had been remarked by many of the participants in the scientific debate, and it was one of the factors that helped to pave the way for the theory of epigenesis.<sup>50</sup>

Kant turns next to epigenesis, which he endorses. Apart from this theory having the upper hand in empirical matters, it is preferable from a rational point of view because it ascribes to nature the capacity to produce and not just to develop. It thus minimizes appeal to the supernatural (CJ 424). According to epigenesis, an organism is a *product*, whereas for preformation it is a mere *educt* (something already formed that is drawn out). As we saw, Kant also calls epigenesis the system of “generic preformation.” It might be somewhat surprising to find epigenesis presented as a kind of preformation theory, as epigenesis and preformation were antagonistic theoretical models. In order to understand Kant’s reason for classing epigenesis under the heading of preformation, it must be recalled that any position committed to *generatio aequivoca* is excluded from the start. As defined in CJ, epigenesis is a variety of *generatio univoca*, and thus it does not attempt to explain organized beings from non-organic matter; on the contrary, it claims that an

---

<sup>49</sup> Pluhar’s translation, which is more accurate here than that of Guyer and Matthews.

<sup>50</sup> See McLaughlin 1990, 17–18.

irreducible teleological factor is indispensable at least at some point in the chain of explanation. In a theory of epigenesis thus conceived, it is assumed that there is a potential for variation in the species that is preformed *virtualiter*, not, to be sure, as a pre-existent minuscule organism, as in the theory of individual preformation, but as a predisposition pertaining to the stock (*Stamme*) (CJ 423).<sup>51</sup> The element of preformation in epigenesis is merely a capacity for a range of possible adaptive variations inherent to a species. In his 1785 review of Herder's *Ideen zur Philosophie der Geschichte der Menschheit*, Kant points out that the "natural destination of formative nature" (*Naturbestimmung der bildenden Natur*) could be designated as "germs or original dispositions" (*Keime oder ursprüngliche Anlagen*), taking these terms to refer not to the implanted buds (*Knospen*) postulated by the preformationist theory of evolution, but merely to limitations (not further explainable) on a self-forming capacity (itself likewise not further explainable).<sup>52</sup>

Thus Kant finds that his favoured version of epigenesis accords with the notion of a formative drive (*Bildungstrieb*) launched by the anatomist and zoologist J. F. Blumenbach. Blumenbach connects epigenesis to the formative drive which he sees as responsible for the directionality in such phenomena as embryogenesis and regeneration of injured body parts. Kant commends Blumenbach on the one hand for his refusal to base life<sup>53</sup> on lifeless matter, and on the other hand for his stating that the formative drive is inscrutable (CJ 424).<sup>54</sup> For Blumenbach the drive is inscrutable in regard to its ultimate nature, but knowable through its effects, just as other basic forces are

---

<sup>51</sup> The concept of an original stock containing purposive dispositions to variation was developed in Kant's earlier essays on human races. See for instance Ak. 8:177. As all functional variation is conceived as originally contained (though not necessarily as material parts) in the original dispositions of the species, this theory is very different from the conception of variation in modern biology.

<sup>52</sup> Ak. 8:62-63.

<sup>53</sup> Kant uncharacteristically uses this term in referring to Blumenbach's theory in CJ 424. Normally he prefers "organization" in biological contexts, reserving "life" for self-determination involving the will.

<sup>54</sup> Though apparently Blumenbach conceived of the formative drive as able to organize unorganized matter, which is incompatible with Kant's reconstruction of the notion. See Richards 2000, 30.

described in the Newtonian tradition. Kant accepts the concept of a formative drive as a heuristic tool for the study of organisms, but in virtue of its being a teleological principle he accords it a merely regulative status. The assumption of an irreducible organization made in the theory of epigenesis must be seen negatively, as a denial that a complete mechanistic explanation can be reached, rather than as a positive appeal to the supernatural. This is presumably the sense of a remark Kant reportedly made in a lecture: “The system of epigenesis does not explain the origin of the human body, but rather says that we know nothing of it.”<sup>55</sup> Seen in this way epigenesis fits with the regulative approach to teleology characteristic of Kant’s third *Critique*.

It should be noted that not all positions customarily referred to as epigenetical belong to that class as it is defined in CJ § 81. For instance, Buffon had proposed an epigenetic theory based on forces acting on organic particles. As spontaneous generation (which is a form of *generatio aequivoca*) apparently figured in his account,<sup>56</sup> it is not epigenesis in Kant’s restricted use of the term. This helps to explain the distinction in *Reflexion* 4851 (1770s) mentioned above, according to which epigenesis can be of two kinds: *a posteriori* and *a priori*. Though the context here is epistemological, the analogy is from the theory of generation. The concept of epigenesis used is the broad one, including epigenetical theories of the *aequivoca* type excluded in CJ § 81. The passage on the “epigenesis of pure reason” (B166–168) shows that Kant in the second edition of the *Critique of Pure Reason* (1787) had adopted the stricter definition of epigenesis, which excludes *generatio aequivoca*. That Kant earlier used the more inclusive notion can also be seen in *The Only Possible Argument in Support of a Demonstration of the Existence of God* (1763), which contains some remarks on the theory of generation which point forward to the discussion in CJ. Without explicit mention of these terms, preformation is distinguished from epigenesis. Buffon’s view is counted to the latter group, and no separate class of *aequivoca* theories is listed.<sup>57</sup> The distinction turns on whether the formation of

---

<sup>55</sup> *Metaphysik K<sub>2</sub>* (early 1790s), Ak. 28:761 (Kant 1997b, 402). Compare Zammito 2003, 88.

<sup>56</sup> See Bäumer 1996, 323–324.

<sup>57</sup> Ak. 2:114–115.

organisms is due to divine action, or based on a capacity on their part to generate their kind.<sup>58</sup> A corresponding terminological distinction is made between “develop” (*entwickeln*) and “generate” (*erzeugen*),<sup>59</sup> where development is the unfolding of a preformed being, whereas generation is epigenetic production. Though the theories of generation of Buffon and Maupertuis “are either as incomprehensible as the thing itself, or [...] are entirely arbitrary inventions”, the failure of these attempts should not force one to opt for a supernatural explanation.<sup>60</sup> Even though epigenesis is not mentioned, it is clear that Kant has in mind the broad notion that includes mechanistic variants such as that of Buffon.

In the analogy in CPR (B167) between epigenesis and the categories as grounds of the possibility of experience, we find the restricted notion of epigenesis that excludes *generatio aequivoca*, the one that is later explicated in CJ. It is now time briefly to look again at the analogy in the light of the distinctions of CJ § 81. A parallel is drawn between *generatio aequivoca* and an empiricist explanation of the *a priori*. In both cases something external is appealed to as explanation (the mechanism of matter for organisms, and sensations for *a priori* concepts), and in both cases the explanation is not adequate to the thing to be explained. The view according to which the categories are implanted predispositions for judgments is compared to the system of preformation (B167). This corresponds to what in CJ is called individual preformation. The flaw of the biological theory is that in its appeal to theological speculation it departs from the task of explanation. The corresponding fault is committed in the epistemological case. Specifically, Kant’s objections to the postulated harmony of categories and experience are, firstly, that there would then be no way of knowing if a concept is a category, since any concept could be a divine predisposition, and secondly, that the relation between a predisposition to think in a certain way and what actually is the case would be entirely contingent, which goes against the necessity pertaining to categories (B167–168). There is no strict parallel between these objections and the empirical weaknesses of the biological theory of preformation.

---

<sup>58</sup> Ak. 2:115.

<sup>59</sup> Ak. 2:115.

<sup>60</sup> Ak. 2:115.

But one can see a similarity on a general level in that the epistemological and the biological theories of preformation both suffer from an irremediable tension between supernatural and natural occurrences. In the epistemological case, either the harmony is complete, but then (absurdly) no concepts are ever wrongly applied, and the distinction between empirical and *a priori* from which we began vanishes; or experience is more modestly held to be independent of the innate dispositions, but then these have no necessary agreement with experience, which leaves their role quite unclear. In the biological case, preformation either collapses into a wholly supernatural occasionalism, so that we “fall into a complete hyperphysics, which could dispense with all natural explanation” (CJ 423), or, if the more modest alternative of preformed germs placed in an independently active environment is chosen, the connection is broken between that which is preformed and the actual outcomes of generation processes as shown in experience, leaving phenomena like hybrids unexplainable.

Epistemological epigenesis considers categories to be “self-thought *a priori* first principles” (B167). Likewise, in biology, the organism is thought of as a “self-organizing being” (CJ 374). Epigenesis is the theory according to which nature has the capacity to produce (*erzeugen*), a capacity it is denied by the theory of preformation. Epigenesis in the narrow (teleological) sense brings generation into the realm of nature without falling into the mechanistic view of *generatio aequivoca*.

In the biological concept of epigenesis as Kant delineates it in CJ § 81, the feature that makes it possible to call it “generic” preformation, namely that it involves a formative drive with predispositions which are “preformed *virtualiter*”, closely parallels the peculiarity of the *acquisitio originaria* theme: that spontaneity is viewed as “innate” ground (corresponding to the formative drive) but no innate representations (and correspondingly, no preformed germs) are admitted. In view of these similarities between the biological and the epistemological cases, it is not hard to appreciate that Kant found the analogy to be of some use.

It must be stressed, though, that the comparison of transcendental philosophy and biological epigenesis is an analogy.<sup>61</sup> That is, in no

---

<sup>61</sup> Compare Zöllner 1988, Ingensiep 1994, and Haffner 1997, 160–172.

way does it imply that the epistemological elements are to be explained biologically. Kant's view is very clear: *a priori* cognition cannot be derived from empirical sources. This also puts some constraints on attempts to use the development of Kant's views on generation to trace changes in his epistemological views, a theme which will be returned to in the next section. Nobody believes that transcendental idealism is based on geography just because CPR contains geographical analogies, and likewise it should be clear that an epigenetical account of the categories is not part of the biological theory of epigenesis. A contingent occurrence, such as the epigenetic process of development, pertains to the empirical realm, and on Kant's view it cannot therefore ground the *a priori* level with its concomitant necessity. Any attempt to found conditions for the possibility of empirical cognition on some instance of this same empirical cognition fails to attain such necessity. In the concluding remarks below it is suggested that apart from this general objection, there is a further reason why biology as conceived by Kant is even less suitable a candidate for such a grounding of philosophy than other empirical sciences would be. But before turning to that I shall take a critical look at a recent attempt to take the analogy to biological epigenesis as more than an analogy.

### 3. A BIOLOGICAL INTERPRETATION OF THE *A PRIORI*

In a recent article, Phillip Sloan (2002) carefully traces the development of Kant's views on epigenesis. He argues that a moderate preformationism was replaced by a moderately epigenetic position in the middle of the 1780s. In a first part of the present section, I shall look at some aspects of Sloan's reconstruction of this development, in so far as it concerns biology. In the second part of this section, what I take as a serious problem with Sloan's approach is examined. The problem is that Sloan seems to take for granted that any mention by Kant of preformation or epigenesis, in any context, philosophical or biological, always at bottom refers to the biological realm, an assumption that leads him to regard the categories as

biological predispositions. The translatability of philosophical contexts to biological theory is not so much something argued for, as rather a presupposition structuring his entire interpretation. In view of the radical discrepancy between an *a posteriori* grounding of transcendental idealism on empirical facts and the explicit claims of necessity made in this philosophy, such a presupposition needs to be provided some justification. Lacking that, the alternative view, according to which the theme of epigenesis has an illustrative function in epistemological contexts, should be the more plausible interpretation, as it fits with Kant's claims concerning the *a priori* status of his undertaking.

### 1) THE DEVELOPMENT OF KANT'S VIEWS ON DEVELOPMENT

Sloan offers a useful overview of the development of Kant's thinking concerning the theory of generation, relating it to the debate in the life sciences of the eighteenth century. From his account I extract the following stages in Kant's theorizing, which I will subsequently comment.

1) An early interest for the question of generation is shown in the 1763 treatise *The Only Possible Argument in Support of a Demonstration of the Existence of God*, in which a compromise is sought between explanations for generation that appeal to the supernatural, and such that accord plants and animals a natural capacity to procreate. Kant favours the latter alternative, but does not find the proposals of Buffon and Maupertuis convincing, as they are too mechanistic to do justice to the teleological aspects of organisms.<sup>62</sup>

2) A technical distinction between *Keime* (germs) and *Anlagen* (predispositions) is introduced in Kant's first essay on the human races (1775). Kant states:

The foundations which lie in the nature of an organic body (plant or animal) for a determinate development [*bestimmten Auswickelung*] are called germs [*Keime*] if this development concerns

---

<sup>62</sup> Kant must also have felt uneasy about Maupertuis' attribution of memory and desire to living particles. See Jacob 1993, 79.

## Acquisitio Originaria and Epigenesis

particular parts; but if it concerns only the size or the relations of the parts to one another, I call them natural predispositions [*Anlagen*]. [...] Chance [*Zufall*] or general mechanical laws cannot bring forth such mutual fit [*Zusammenpassungen*]. Therefore we must regard such occasional developments [*Auswickelungen*] as preformed [*vorgebildet*]. But even where nothing purposive is displayed, the mere capacity to propagate its particular acquired character is already sufficient proof that a particular germ or natural predisposition for it must have been present in the organic creature.<sup>63</sup>

Sloan stresses the distinction between *Keime* and *Anlagen*, which he construes as pre-existent germs (*Keime*) underlying the organism's traits, acted upon by a structuring power or disposition (*Anlagen*).<sup>64</sup>

3) A next phase is marked by Kant's first use in his published writings of the term "epigenesis," in his 1785 review of Herder's *Ideen*. Though generally negative towards what he viewed as the uncritical and speculative excesses in Herder's vision of the unity of human and natural history, Kant agreed to some extent with Herder's postulation of a "genetic force" (*genetische Kraft*). As Kant explains, with this Herder "wants on the one hand to reject the system of evolution, but on the other also the purely mechanical influence of external causes, as [both being] useless explanatory grounds."<sup>65</sup> In *The Only Possible Argument* (1763), Kant had also looked for a third alternative, between preformation (that is, the system of evolution) and mechanism. He thus agrees with Herder on this issue, but with a proviso:

if the cause organizing from within would be by its nature limited only to perhaps a certain number and degree of different formations of the creature [*Geschöpfs*] (and if after the execution thereof it would not be free to form another type in changed circumstances), then one could just as well designate this natural

---

<sup>63</sup> *Von den verschiedenen Racen der Menschen* (1775), Ak. 2:434–435. The translation is based on those given in Sloan 2002, 240, McFarland 1970, 58, and McLaughlin 1990, 31.

<sup>64</sup> Sloan 2002, 239–240.

<sup>65</sup> Ak. 8:62, Sloan's translation (2002, 243) modified.



destination of formative nature germs or original predispositions, without therefore regarding the former as machines<sup>66</sup> and buds (as in the system of evolution) that are enclosed from the first beginning and only on occasion unfold themselves, but rather as bare, not further explicable limitations on a self-forming capacity, which latter we can explain or make comprehensible just as little.<sup>67</sup>

Kant's proviso expresses a restriction on how far epigenesis can be allowed to reach, as well as a clarification of his use of "*Keime*" and "*Anlagen*." The restriction states that one should consider the number of preadaptations inherent in a species to be fixed, and that such (functional) variation, once it has taken place, cannot be followed by yet more variation in new circumstances, at least not after a certain point in the teleological development. This restriction is central in Kant's essays on races and serves, anachronistically speaking, to prevent an evolutionary view (in the modern sense of the word) on the transformation of species. The clarification of the concepts "germ" and "predisposition" amounts to an excision of all elements pertaining to stronger theories of preformation (such as the notion of encapsulated buds). According to Kant, a teleological theory of epigenesis can speak of germs and predispositions without assuming more than a self-forming capacity with inherent limitations.

Sloan comments on Kant's reaction to Herder's view in somewhat different terms, but presumably with the same intent: *Keime* and *Anlagen* "have assumed the role of limiting structures on the *Lebensprincip*."<sup>68</sup>

4) A further development in Kant's theorizing on generation reflects the impact of J. F. Blumenbach's epigenetic theory of the formative drive (*Bildungstrieb*). Though Kant mentions Blumenbach earlier, it seems that it was the second edition of Blumenbach's *Über den Bildungstrieb* (sent to him by the author in 1789) that occasioned

---

<sup>66</sup> In the preformation theory of Leibniz, organic bodies were called "machines." See *Monadology* (1714) §§ 64 and 74 (Leibniz 1951, 546 and 548).

<sup>67</sup> *Recension von Herders Ideen, Zweiter Theil*, Ak. 8:62–63. Translation based on Sloan 2002, 243–244, but heavily modified.

<sup>68</sup> Sloan 2002, 244.

a deeper involvement in the theory of epigenesis.<sup>69</sup> Blumenbach, a respected scientist, developed epigenesis into a more sophisticated position, adducing biological evidence for the presence in organisms of a general formative drive, detectable in processes of generation, nutrition and regeneration.<sup>70</sup> This postulated drive (conceived as intrinsically inscrutable but observable via its effects) was linked to a complete rejection of preformed germs. On Blumenbach's view, *Keime* were anyhow unable to explain the facts of generation, and once a formative drive was assumed they were not needed at all. Moreover, microscopic observation failed to give evidence of preformed germs, whereas hybrids and phenomena of regeneration were conformable to epigenesis.

Blumenbach's influence apparently prompted Kant to give up on using "*Keime*" even in the more epigenetic sense of the term explained in the Herder review. As Sloan points out, the word is absent from the third *Critique* (1790),<sup>71</sup> which endorses Blumenbach's version of epigenesis.

These are the main stages, as traced by Sloan, of Kant's development with regard to the strictly biological sense of the theory of generation. Some comments will now be made concerning this development, before turning to the epistemological analogy.

1) Sloan describes how Kant in the first stage (marked by the 1763 treatise on the existence of God) points out the need for a middle view between strong preformation theories on the one hand and Buffon's and Maupertuis' theories on the other. But it should also be noted that Kant undoubtedly leans more towards the general orientation of the latter theorists.<sup>72</sup> For even though their particular views are "incomprehensible" or "arbitrary," at least they accord to organisms a capacity "truly to generate" (*erzeugen*), which is what the theory of preformation denies.<sup>73</sup> So Kant draws near to epigenesis

---

<sup>69</sup> Sloan 2002, 247.

<sup>70</sup> Sloan (2002, 247) cites Blumenbach to this effect, but then a bit misleadingly claims that the formative drive "was restricted to purposive embryological formation" (2002, 248).

<sup>71</sup> Sloan 2002, 248.

<sup>72</sup> Compare McLaughlin 1990, 27: "Kant clearly takes the side of epigenesis in this work."

<sup>73</sup> Ak. 2:115.

already here. Sloan apparently takes Kant's compromise to be rather close to the moderately preformationist germ theories of Albrecht von Haller and Charles Bonnet, distinct from older more extreme versions of the theory.<sup>74</sup> Even on Sloan's account, however, Bonnet's theory "verged on a literal preformation of the embryo in the *germe*."<sup>75</sup> In fact, it seems not only to verge on, but to assert literal preformation of the embryo in the germ. Sloan cites Bonnet in 1762 stating that the germs are "in miniature a man, a horse, a bull, etc."<sup>76</sup> On Haller's view, what is preformed is not simply a miniature embryo, but rather all the parts of the adult organism, though in very different relations and proportions. Hence the unfolding amounts not just to growth, but to a complex adjustment of the growing parts. This theory, like Bonnet's, undoubtedly represents a more sophisticated account than the traditional idea of a little animal waiting inside the germ. Nevertheless, according to Haller "all the essential parts of the fetus exist formed at all times."<sup>77</sup> Thus, on this theory, nothing is really generated anew in reproduction. The Haller-Bonnet germ theory can hardly meet Kant's 1763 desideratum on a theory of generation, namely, that it should accord to organisms the capacity to generate, not just to develop.<sup>78</sup>

2) The second phase in Kant's development begins in 1775 with the technical distinction between *Keime* and *Anlagen*. Here Sloan construes *Keime* as pre-existent germs underlying the parts of the organism, and *Anlagen* as a structuring power that coordinates the development of these parts. He views this power as a teleological predisposition that adjusts the parts "to one another, and to their

---

<sup>74</sup> Compare Sloan 2002, 235–236.

<sup>75</sup> Sloan 2002, 236.

<sup>76</sup> Sloan 2002, 236. Bonnet in this passage stresses that the preformed miniatures have the general traits of the species but not of the individual. In this way the environment gets a role in explaining variety, but it remains a strongly preformationist view. A complete organism is enfolded in the germ.

<sup>77</sup> Haller as quoted by Sloan 2002, 236.

<sup>78</sup> Ak. 2:115. Kant has *entwickeln* here, as it seems synonymously to *auswickeln* used before in the same passage. This speaks against Sloan's claim (2002, 240n 38) that the former term has less of a preformationist sense. (Though this might be the case in some other writers of the period).

surrounding conditions,” bringing a new element of environmentalism to the theory of preformation.<sup>79</sup>

Kant’s distinction, however, is more narrowly conceived. “Germs” designate the foundations for particular parts, “predispositions” the foundations for the size or mutual relations of these parts. Kant gives the example of a species of birds with germs for developing an extra layer of feathers in cold climates which are not developed when the birds live in warmer climates. Wheat grains getting a thicker hull in cold countries is his example of a predisposition.<sup>80</sup> The distinction seems to rest on viewing the thickening of the hull as an increase in size, without any extra part added, whereas the new layer of feathers is a separate structure, which therefore needs an underlying germ. Lacking in Kant’s account is the interplay suggested by Sloan between germs as parts and predispositions as teleologically arranging these parts according to what the environment demands. Rather, germs and predispositions are both thought to act teleologically, they both develop in response to environmental circumstances. The distinction only concerns whether the teleological factor produces a separate part or an alteration of size or relation with regard to parts, without suggesting a special regulating role with respect to the environment for the predispositions. If a germ brings about the growth of an extra layer of feathers for the sake of coping with the cold climate in the bird’s habitat, predispositions take part in this process only in so far as the new feathers must affect parts of the previous body type. Sloan’s construal of *Anlagen* as foundations for the relationship of parts “to their surrounding conditions” has no counterpart in the text he refers to (Ak 2:434–435). Moreover, it is not clear that Kant sticks to the distinction between germs and predispositions in any strict way in other texts. This brings us over to the third stage.

3) In his comment on Herder, Kant uses the terms together. But the distinction gets blurred. Kant states that on admitting Herder’s inner “genetic force,” it should be all right to call it “*germs or original predispositions*,” provided that these terms are not taken in the preformationist sense. On a rather natural interpretation of the passage, it seems to say that it matters little whether we call the inner

---

<sup>79</sup> Sloan 2002, 239–241.

<sup>80</sup> Ak. 2:434.

force a predisposition or a germ, as long as we are not postulating any preformationist “buds” but merely refer to an ultimately inexplicable capacity. In later texts one finds the terms used interchangeably.<sup>81</sup>

As Sloan notes, under the influence of Blumenbach Kant drops *Keime* altogether in stage 4). Now, if the present reading is correct, he could have continued to use the term, given that its epigenetic translation proposed in the review of Herder took away its preformationist content. But after Blumenbach’s attack on *Keime*, Kant probably felt that continued use of the term was bound to be mistaken for a commitment to the more traditional position, and that the word *Anlage* had the more appropriate connotation. Having long since ceased to apply the technical distinction between germs and predispositions, he had no reason to retain a potentially misleading word.

The picture that emerges shows a moderately epigenetic Kant during the whole period from 1763 to 1790, rather than the moderate preformationist turning into a moderate epigeneticist depicted by Sloan. Especially the early phase and the late one resemble each other. The passage in CJ § 81 delineating the alternatives available for the theory of generation has much in common with the text from *The Only Possible Argument*. Both texts stress the importance of giving an explanation as parsimonious as possible with respect to the number of supernatural occurrences to be assumed (without falling into *generatio aequivoca*), a criterion that somewhat unexpectedly makes the occasionalist doctrine of direct divine action at the moment of each fertilization methodologically sounder than preformationism. Also, both discussions focus on the fundamental difference between theories attributing to nature a capacity genuinely to produce and theories claiming that nature only unfolds pre-existent beings (*product* versus *educt* in the CJ terminology), favouring the former. Perhaps there are better grounds for ascribing a more preformationist view to the Kant of 1775, who formulates technical definitions for *Keime* and *Anlagen* and considers all adaptive changes that a species may undergo to be pre-given potentially in its original “stock.” In a

---

<sup>81</sup> See for instance Ak. 8:172–173.

loose, general sense, this is as preformationist as one can get.<sup>82</sup> In a narrower technical sense, however, it is not necessarily preformationism. It is teleology, even naïve teleology, not untypical for German Enlightenment philosophy (though original in its details), but it is not preformation in the narrow sense if it does not assume pre-existent germs actually (not just virtually) containing all parts to be unfolded. Whether Kant in 1775 accepts the Haller-Bonnet view's actual preformation of the parts is not easy to determine.<sup>83</sup> Ten years later, in his review of Herder, he opts for an epigenetic construal of the germs and predispositions, viewing them as modifications of a capacity. This conception is unchanged in the third *Critique*, where the removal of the word *Keime* is of little consequence for the rest of the theory, and epigenesis is classed as *generic* preformation containing predispositions *virtualiter*.<sup>84</sup> One circumstance that might favour a modestly epigenetical reading of Kant's 1775 view is that the germs so suggestive of preformation theory are said to belong to the capacity to generate, the *Zeugungskraft*.<sup>85</sup> This word sounds very epigenetic. But then again, a preformationist might use such a word in the context of *evolutio*, rather like how Kant in 1785 proposed to use *Keime* in the theory of epigenesis. Perhaps a more important indication of an epigenetic orientation is that Kant takes for granted in the 1775 essay on human races that phenotypic traits like skin colour are influenced by both parents – a view that is incompatible at least with strong preformationism, which does not accord the parents any real capacity to generate. On the whole, I think the general outlook in Kant's essay is rather empirical (though in a teleological framework); germs and predispositions are posited as theoretical terms without commitment to any explicit hypothesis about their nature, rather like how genes were posited in early twentieth century genetics to account

---

<sup>82</sup> In this general sense, even the archepigeneticist Herder, because of the teleological orientation of the cosmic development he envisaged, can be said to represent a “preformist vitalism” (Palti 1999, 333).

<sup>83</sup> Haffner (1997, 128) suggests that Kant himself was uncertain at the time.

<sup>84</sup> This is not to deny that the theory is placed in a new light with the introduction of the reflective power of judgment in CJ. But the changes concern the meta level of philosophical architectonic more than the object level content of the biological and anthropological theories.

<sup>85</sup> Compare Ak. 2:435–436.

for the heredity of observable traits, without there being any way to observe the underlying biochemical processes in the organism.

In light of the preceding discussion, it is interesting to note that an early critic of Blumenbach's theory of the formative drive, the physiologist Unzer, in his incisive review of Blumenbach's book pointed out that even on this theory it must be presupposed that there is something on which the formative drive acts, and that this something must be thought to have some special properties making it possible for it to get organized in the appropriate way. Now, according to Unzer, this is just what a sophisticated preformationist would call *Keime*.<sup>86</sup> Though defending the theory of preformation against its new rival, Unzer makes a move somewhat akin to Kant's annexation of *Keime* to the theory of epigenesis in 1785. In both cases, it is claimed that a formative drive must be subject to some constraints. There is no reason to assume that Kant had seen Unzer's review, though it is quite possible. In any case, in the late 1770s, while composing the *Critique of Pure Reason*, he had carefully read J. N. Tetens' *Philosophische Versuche über die menschliche Natur und ihre Entwicklung* (1777),<sup>87</sup> and this book contains a long discussion on *evolutio* versus epigenesis, reaching a compromise not so far from either Kant's or Unzer's later views.<sup>88</sup> Tetens' compromise was put as a slogan: "evolution through epigenesis," meaning

---

<sup>86</sup> Unzer 1782, 167.

<sup>87</sup> Haffner (1997, ch. 3) looks into other possible sources for Kant's knowledge of the concept of epigenesis, such as a book by Erxleben from 1768. That Kant early on was familiar with the term and also connected it to his epistemological views is shown by its occurrence in several *Reflexionen* estimated to be from the 1770s (see for instance R4104, Ak. 17:416; R4275, Ak. 17:492; R4552, Ak. 17:591; R5462, Ak. 18:189; and the already cited R4851). The dating of Kant's notes is not beyond doubt, but Zammito's (2003, 85n) suggestion that these are from a much later period seems to be based less on evidence than on the wish to bolster his hobby horse, Kant's "stealing" epigenesis from Herder (see Zammito 2003, 93). If this is supposed to mean merely that Kant, because of or in spite of his rivalry to Herder, might have been influenced by reading him into making more of the analogy between philosophical views and epigenesis, it might well be true.

<sup>88</sup> Tetens has almost 100 pages on the theory of generation, with comparisons between Bonnet's preformationism and K. F. Wolff's epigenetic theory. See Tetens 1979 [1777], 448–538.

that there is truth on both sides, since epigenesis presupposes some structure for its formative capacity to work on (not just anything gets formed into an animal), while *evolutio* must accept that in its unfolding of the pre-given structure many new structures are formed epigenetically.

Sloan's scheme on the one hand and my defence of continuity on the other give two alternative pictures of Kant's development. They are not dramatically different, but the difference is of some importance when we proceed to the comparison to Kant's employment of biological analogies in philosophy.

## II) ARE THE CATEGORIES GERMS?

In the above account biology is separated from philosophical uses of concepts of the theory of generation in Kant. Sloan's point, however, is that no such separation should be made. His thesis is that the categories literally are biological germs and predispositions, and that phrases in Kant's *Critique of Pure Reason* normally read as analogies are part of a general biological theory.

Support for this interpretation Sloan finds in the changes over time in Kant's use of biological language in central philosophical contexts. As these changes are correlated with the stages of Kant's changing biological views as traced by Sloan, he finds it compelling to see the philosophical tenets, especially those which concern the status of the categories, as not just parallel to tenets in generation theory, but as actually belonging to it. If correct, such an interpretation not only helps in explaining difficult issues about the categories, something that also might be accomplished by a consideration of the parallel to generation theory which takes it as merely an analogy; it could actually provide a lot of new information. For if the theory of categories is not just *like* a system of epigenesis (as a biological theory) in *some* respects, as the proponent of an analogical reading holds, but truly *is* such a system, then we can bring many more facts about biological epigenesis (as understood by Kant) to our interpretation of the categories, gaining much knowledge about them unattainable from Kant's philosophical texts.

The most important examples of the use of terminology from the theory of generation in strictly philosophical contexts in Kant are in



CPR A66/B90–91 and B167. The latter is the passage comparing the categories to a system of the epigenesis of reason. The former one, from the first edition of CPR in 1781, has a decidedly different, preformationist tone. Here, Kant explains that his method is no mere conceptual analysis, but an analysis of the faculty of understanding, aiming to find the *a priori* concepts.

We will therefore pursue the pure concepts into their first seeds and predispositions [*Keimen und Anlagen*] in the human understanding, where they lie ready, until with the opportunity of experience, they are finally developed [*entwickelt*] and exhibited in their clarity by the very same understanding, liberated from the empirical conditions attaching to them. (A66)

According to Sloan, we should consider the use of *Keimen und Anlagen* here as technical, following the definitions given in 1775. Thus, the innate germs are “preformed causal agencies,” whose development is coordinated by *Anlagen*.<sup>89</sup> The mediating role of the predispositions makes it possible for the categories as “biologically pre-existent” germs to respond to external conditions.<sup>90</sup>

As argued above, the predispositions do not really have the role of coordinating the germs to the environment in Kant’s biological explication of this terminology. So the special, more environmentally responsive form of innatism Sloan finds here has no textual basis even on a biological reading of the terms. Moreover, as it is quite possible that germs and predispositions were conceived of in a moderately epigenetic sense even in 1775, a biological interpretation of CPR would have to take this possibility into account. But on a more fundamental level, the equation of the categories to biological structures entirely disfigures their *a priori* status. In view of Kant’s strong claims on finding *a priori* conditions for experience, there is no reason to construe the terms in a literally biological sense, irrespective of whether one prefers a preformationist or an epigenetic model.

Kant’s introduction of epigenesis as a model for understanding the categories in the B-edition of 1787 is likewise interpreted biologically

---

<sup>89</sup> Sloan 2002, 241.

<sup>90</sup> Sloan 2002, 242.

by Sloan. He links the shift from preformation to epigenesis to Kant's reading of Herder in 1785, which led to a new conceptualization of germs and predispositions in terms of a self-forming capacity rather than as preformed "machines." This is plausible, if taken in a rather limited sense. Upon reading Herder, Kant turned a bit more epigenetic in his biological views (and as the above discussion may have shown, it was perhaps no large step: nothing indicates that he ever believed in the preformationist's "buds" and "machines"). This greater interest for the biological theory of epigenesis, as well as for Herder's use of it in philosophy, led him to adopt it as a metaphor for explaining the status of the categories. As Günter Zöllner has argued, epigenesis was an appropriate model for explaining the understanding's generation of *a priori* concepts.<sup>91</sup> Generation of concepts taken as an active process, which is also found in the metaphor of *acquisitio originaria*, is better conveyed in the epigeneticist's terms of self-forming capacities than in the more passive language of preformed structures occasionally used in the A-edition.

Sloan, however, goes much further than pointing to a shift in metaphors. According to him, the categories are "biological properties."<sup>92</sup> This interpretation would require some revisions of standard views on Kant's notion of the *a priori*. It is hard to see how the categories could be necessary in any strong sense if taken as biological, and Kant always emphasizes their necessity in quite uncompromising terms.

On the interpretation that takes epistemic uses of generation theory as metaphors, it is easier to explain Kant's seemingly inconsistent retention of the preformationist A66 passage in the B-edition. After all, a metaphor should be able to survive the encounter with its opposite metaphor better than is the case with opposed theories. Epigenesis nicely illustrates the understanding's self-generating activity in bringing forth the *a priori* conditions of experience, but nonetheless, if the analogy is not taken too far, there are also features of the idea of preformation that can be of use. The image in A66 of the categories lying ready, untouched by anything empirical, effectively puts across the notion of the complete separation of the *a priori* from the empirical. The word *Keime* also has the sense of

---

<sup>91</sup> See Zöllner 1988, 88.

<sup>92</sup> Sloan 2002, 245.

“seeds,” which is a rather commonplace metaphor easy to use regardless of one’s preferences in generation theory.<sup>93</sup> But preformationist language has the inconvenience of being associated with innatism. It is interesting in this context to note that Brigitte Sassen considers the insertion of the section on epigenesis in the B-edition to be Kant’s reaction to his friend Johann Schultz’s suggestion that the categories might have their basis in a preestablished harmony.<sup>94</sup> Though this may not be a full explanation, it is probable that Kant’s main reason for introducing epigenesis in the critical philosophy was to counteract interpretations of it in terms of innatism. This he could have done even if he had not believed in epigenesis as a biological theory. The constraints on the adoption of metaphors are certainly quite fluid. Kant would presumably not have chosen epigenesis as analogy for the function of the categories if he had deemed it worthless as a biological theory; but as long as a theory is respectable and has an illustrative value there seems to be no reason not to use it as metaphor, regardless of one’s view on its validity in its own domain. That is one of the advantages of metaphors: you choose them for their illustrative power rather than for their truth. So much the better if you also think the theory is valid, as Kant appears to have thought most of the time of epigenesis in biology.

#### 4. CONCLUDING REMARKS: REASON AND THE ORGANISM

The most prominent of Kant’s biological analogies in philosophy is that between reason and the organism. There is a structural isomorphy between the concepts of reason, system, and organization. Reason is a system (CPR A738/B766), and a system is articulated “like an animal body” (A833/B861). Reason’s activity is also seen in such terms. In its historical development, the early attempts at system building appear as if the systems had been formed “by a *generatio aequivoca*

---

<sup>93</sup> It is therefore not necessary to follow Sloan (2002, 250) in taking Kant’s statements on *Keim des Guten* in ethical and theological contexts to refer to the biological domain.

<sup>94</sup> Sassen 2000, 301.

from the mere confluence of aggregated concepts,” but they all had their “schema, as the original seed [*Keim*], in the mere self-development of reason [*sich bloss auswickelnden Vernunft*]” (A835/B863). In pure speculative reason, “as in an organized body, every part exists for the sake of all the others as all the others exist for its sake” (Bxxiii). It “contains a truly articulated structure of members in which each thing is an organ” (Bxxxvii). These descriptions parallel the characterization in CJ of an organized being as a “product of nature” in which “everything is an end and reciprocally a means as well” (CJ 376), so that “each part is conceived as if it exists only through all the others, thus as if existing for the sake of the others and on account of the whole” (CJ 373), as “an organized and self-organizing being” (CJ 374).

Many questions could be posed about the relation between reason and the organism, but in the present context I shall only discuss whether the parallel could be made use of for a biological interpretation of transcendental philosophy. After all, it seems that such a close connection should have some significance. What I shall suggest is that rather than opening for a naturalization of reason, the Kantian view is that this parallel comes about through the analogical employment of reason’s own features in conceiving of natural objects.

A first point to note is that the parallel is not complete: the organism is a “product of nature,” reason on the other hand is not bound to the conditions for cognition of nature but transcends them (though without thereby attaining cognition) (A308–309/B365–366). As a particular empirical object, the organism is an object for the reflecting power of judgment, which searches for a universal or rule under which to subsume the particular.<sup>95</sup> Leaving all details aside on the intricacies of the distinction between the reflecting and the determining power of judgment, it is clear that since reason is no object of experience, it cannot be the case that the ascription of systematicity and organization to it is due to the reflecting power of judgment. Whether Kant holds that reason’s own self-knowledge teaches us of its structure, or whether our conception of what reason is like is based on analogy is a difficult question which I leave aside.<sup>96</sup>

---

<sup>95</sup> See CJ 179 and Ak. 20:211

<sup>96</sup> See Kleingeld 1998 for an interesting attempt to cast some light on this.

Now, the organism is not merely judged reflectively as any empirical object for which a concept is to be found, but “we ascribe to it our concept of an end [*Zweck*]” (CJ 193). Thereby we go beyond the general assumption of the purposiveness of nature for our conceptual capacity which defines the reflecting power of judgment, and by analogy represent the object as a natural end (*Naturzweck*), that is, an organism (CJ 193). The use of the analogy of an end depends on reason, which, as faculty of desire, acts “in accordance with the representation of an end” (CJ 220). The self-organizing capacities of organisms prompt us to apply the notion of end analogically for the sake of interpreting them, as we otherwise find the complex arrangement of their parts too difficult to grasp.<sup>97</sup>

In short, biology as the science of the organism has as its horizon the attribution of ends to the objects under study. But as purposiveness cannot ultimately be known objectively to apply to nature also apart from our mode of judging, Kant accords the teleological judgments structuring biology the status of regulativity.<sup>98</sup>

The regulative status of biology puts it in a peculiar position among the natural sciences. Whereas physics is conceived of realistically by Kant (relative to the level of empirical realism), biology is viewed as an interpretive science dependent on analogy. This is a final reason why biological interpretations of Kant’s transcendental philosophy are inadequate. Not only would there be a problem for such an interpretation to explain the status of the *a priori* in terms of the empirical. In addition, it would have to explain how biology, which for Kant is built on an analogy brought to it from reason, could possibly ground that same reason on which it depends.

---

<sup>97</sup> I will return to Kant’s conception of the organism in the following essays.

<sup>98</sup> Though in biological practice, when the meta perspective is dropped, a realistic stance towards the objects under study is assumed. More on this in the next essay.

# Biological Functions in a Kantian Perspective

For the last thirty years or so, the most influential philosophical accounts of biological functions have been framed in terms of etiology and causal role respectively. Each of them has strong points as well as weaknesses. I shall here present these accounts and explain some of the reasons for why the etiological theory, until recently the dominant approach, is currently losing this status to the causal role approach (of which the so called propensity theory can be considered a special form). But the causal role theory is not without its own difficulties. It is argued that there is a problem in delimiting the kind of systems for which causal role analysis should result in functional attributions, and that this calls for introducing standards of biological normality, which are problematical from a purely naturalistic perspective. It appears that we are constrained to presuppose some special principle for regarding natural objects, in the light of which the notion of biological function makes sense.

In order to get a grip on the question of biological normality, it is illuminating to go back to Kant's discussion of the role of the teleological maxim for judging organized beings and bring it in contact with the recent debate on functions. One feature of Kant's account is that functional ascriptions are claimed to be regulative and not constitutive, in that they are made from a perspective that is not as fundamental as that in which ordinary causal ascriptions are made. Kant's view is sometimes called projectivistic, implying that teleology is but a metaphor that we project on natural objects. But since Kant treats the teleological maxim as a condition for the possibility of biology, and not just as a heuristic aid for its study, it is hardly appropriate to regard it simply as metaphorical (disregarding

the question of what exactly the distinction between the metaphoric and the literal amounts to).<sup>1</sup> If a teleological point of view is a condition for biology, teleology should rather be seen as a constitutive principle for this science. This may seem to contradict Kant's insistence on the regulativity of teleology, but when a distinction is made between the point of view of the biologist and the meta level perspective of philosophy, it is possible to construe teleology as being constitutive in the biological perspective but merely regulative on the meta level. By making a comparison with Aristotelian analyses of function in terms of what is good for an organism, I shall attempt to make Kant's view somewhat clearer. The Aristotelian conception can to a large extent be accommodated in the Kantian frame, since the notion of the organism as a being that has a good follows naturally from the teleological maxim adopted in the Kantian stance. What distinguishes the Kantian from the Aristotelian view is that the former treats teleology as constitutive for properties which are, on a meta level, only to be assumed regulatively (namely, traits having functions), whereas the latter, making no distinction of levels, treats functions as on a par with any other properties.

## 1. ETIOLOGICAL FUNCTIONS

The etiological account of biological functions was pioneered by Larry Wright (1973). Simplifying slightly, according to Wright's analysis something has a function F just in case it exists (or continues to be where it is) because it does (or results in) F. To say, for instance, that the heart has the function of pumping blood is to say that its presence is due to its having this effect. To avoid the paradoxical air of backwards causation in this, it should be taken as a claim about the etiology of hearts, in terms of selection. Though the analysis does not explicitly include claims about selection, the fact that we can rightly ascribe functions to some things is explained by considering that these things are selected because they have certain effects. For biological cases the item (to use the term traditionally

---

<sup>1</sup> But compare Ratcliffe 2000 and 2001 for an assessment more favourable to Kantian projectivism.

employed in this debate) that has the function is of a kind favoured by natural selection, whereas for artefacts (for instance an artificial heart) the selection is made by a conscious agent. Wright wanted to provide a quite general account of function ascriptions, valid in biology as well as for artefacts. In conjunction with certain vaguenesses in Wright's analysis (for instance with respect to the distinction between type and token),<sup>2</sup> this made it vulnerable to counterexamples.<sup>3</sup>

It was soon perceived that a completely general analysis of function ascriptions was difficult to achieve and perhaps not even desirable. Etiological accounts such as Karen Neander's (1991a; 1991b) and Ruth Millikan's (1989b) aimed only at an explication of biological function, and were therefore free to define the notion directly in terms of natural selection. In Neander's analysis, an item X has a function in an organism O if and only if X does what instances of X's type did that contributed to the fitness of O's ancestors and by virtue of which the genotype expressed in phenotype X was selected.<sup>4</sup>

Connecting the concept of function to natural selection in this way gives it a clear scientific meaning, and the analysis has other advantages as well. For instance, the problem of how to account for the difference between a function and a mere effect is resolved by distinguishing the trait "selected for" (selected in virtue of its fitness) from traits accompanying it but lacking selective significance. The standard example is the heart's pumping the blood (a function) in contrast to its sound (a mere effect); the former is a property that has been selected for, the latter certainly not. Another advantage claimed for the etiological account is its handling of the distinction between function and malfunction. A heart that, because of damage or deformation, does not pump the blood, still has the function of pumping the blood. It is malfunctioning, but it has a function. It is "*supposed to*" pump blood.<sup>5</sup> The etiological account is claimed to

---

<sup>2</sup> See Nissen 1993, 30–31.

<sup>3</sup> See Bedau (1991, 648) for one of many counterexamples to Wright's analysis.

<sup>4</sup> Compare Neander 1991a, 174. Though it is not uncontroversial that natural selection should be viewed as operating only on the genotype level, this issue will not be pursued here.

<sup>5</sup> Neander 1991b, 467.



preserve the normative status of functional ascriptions, whereby an item can have a function that it ought to perform even when it does not do so. This is because a biological category like the heart is defined by its function, and its function is its selected effect.<sup>6</sup> Therefore the heart as type has the function of pumping blood, and any present token heart has that function in virtue of its position in a lineage of hearts once selected for pumping, no matter if it happens to be malfunctioning.

In the further development of the etiological approach, the definition of an item's function as being the feature for which it was originally selected has been challenged by the view that what should be focused on is not ancient history, but the selective pressures in force in more recent evolutionary history.<sup>7</sup> The rationale for this emendation of the theory is that an organ can change its function over time, and in such a case it seems more to the point to consider what was recently selected to be the function, rather than some feature no longer evolutionarily relevant. The usual example here is the original function of feathers in birds. On one hypothesis, they were selected for insulation, and only later was their contribution to effective flight selected for.<sup>8</sup>

A further refinement was made by David Buller (1998) in distinguishing a stronger and a weaker version of the etiological theory. The stronger version requires for a trait of an organism to have a function that it is hereditary and was selected for in the presence of variants doing something similar but contributing less to the fitness of the ancestors concerned. The weaker version accepts a trait as functional if it is hereditary and has contributed to ancestral fitness, even if there was no selection for that trait. The weaker theory can avoid some of the disadvantages of the strong theory. There are for instance empirical studies indicating that there is but a very small amount of variation in basic biochemical traits in wild populations. If variation has not occurred in these cases, the strong etiological theory must, counterintuitively, deny functions to such traits. With a less

---

<sup>6</sup> See Millikan (1989b, 295) and Neander (1991a, 180) on the functional categorization of organs.

<sup>7</sup> Godfrey-Smith 1994.

<sup>8</sup> Godfrey-Smith 1994, 357–358.

demanding definition of function, requiring ancestral fitness but not variation, traits lacking in variation can be accorded functions.<sup>9</sup>

The etiological analysis of function is closely connected to the concept of adaptation.<sup>10</sup> This concept, as defined by Elliott Sober (1993, 84), is essentially equivalent<sup>11</sup> to the strong version of etiological function:

Characteristic *c* is an adaptation for doing task *t* in a population if and only if members of the population now have *c* because, ancestrally, there was selection for having *c* and *c* conferred a fitness advantage because it performed task *t*.

As can be seen in Buller's (1999) survey of analyses of biological functions, until very recently etiology in some version was the received view. This seems to be changing, and I will indicate some points of criticism that motivate alternative approaches. To some extent, the problems that critics have found in etiology can be traced to a tension between the somewhat opposing motivations for the theory. On the one hand, the interdefinability of function and adaptation (where different versions of etiological function correspond to variant concepts of adaptation) would appear to amount to a reduction of teleological locutions to a naturalistically safe concept that is part of current biology and depends only on efficient causation. Such a view is in line with the perception that Darwin purged biology of Aristotelian teleology. On the other hand, it is also claimed that the etiological theory makes function respectable in biology, even as a normative and "genuinely teleological notion."<sup>12</sup> The normative dimension in functional attribution is accounted for by the historical background. A trait's function is defined by what its counterparts in earlier generations were selected for, regardless of whether the trait is able to perform this now.

---

<sup>9</sup> Buller 1998, 512–513.

<sup>10</sup> Melander (1997, 91) is one author who explicitly defines function as adaptation (though he also recognizes another, non-etiological concept of function).

<sup>11</sup> Following Buller's analysis of "selection for" a trait, according to which this notion implies the existence of variations with respect to the trait. See Buller 1998, 508.

<sup>12</sup> Neander 1991b, 467.

The normative claims of the etiological analysis can be challenged. It does not seem evident that something is supposed to do something just because something else did that same thing. An observer will certainly expect that a hitherto unexamined heart is pumping blood if she knows that it belongs to a lineage of organs originally selected because of that capacity (disregarding other reasons for the expectation besides knowledge of selectional history). But is this heart therefore *supposed to* pump blood, or is it the observer that supposes – on good inductive grounds, to be sure – that it will do that?

Furthermore, in an influential paper Amundson and Lauder argues that it is normally not the case that biological items are defined in terms of function.<sup>13</sup> More basic is the relation of homology. Two traits are homologous if they have a common descent; for instance, a human arm and a bat's wing are homologous, since these structures have the same evolutionary origin. Homologous organs can have different functions, so “if anatomical parts had to be identified by their common [...] function, all interesting homologies would be invisible.”<sup>14</sup>

This argument casts some doubt on the historical link whereby a lineage of functioning hearts is supposed to confer a norm to a present, malfunctioning heart, in virtue of their belonging to the same functional category. If organs typically are categorized not functionally, but on the basis of homology, allowing different functions for items in the same category, the etiological approach will not, in most cases, be able to appeal to this historical model to account for the normativity of function. However, in the particular case of the heart, we have a case where an organ *is* functionally defined, which Amundson and Lauder (1994, 455) admit, leaving the possibility that Neander's account is correct at least about hearts.

A more direct argument against the normativity of etiological function is offered by Paul Sheldon Davies (2000a). Suppose that hearts are functionally categorized in terms of success in the performance of a task. If now a present, damaged heart is unable to perform the task, it cannot be a malfunctioning heart. Actually, it can't be a heart at all, because the category “heart” is etiological

---

<sup>13</sup> Amundson and Lauder 1994, 454; see also Wouters 1999, 213.

<sup>14</sup> Amundson and Lauder 1994, 455.

characterized by the successful performance of what selected hearts did. According to a strong etiological theory (as characterized by Buller), hearts must have varied as to fitness. From all these hearts, constituting the generic type, a narrower type was selected for its greater fitness. This narrower type is the selected functional type. But consequently, if a present heart loses the property for which the narrower type was selected, it thereby no longer belongs to the narrow type. It is rather a heart in virtue of belonging to the generic type, in which also the less well-functioning items are included. The etiological theory, however, assigns function precisely to the successful narrow type. Malfunctioning hearts are therefore excluded from the functional (narrow) type, and cannot be counted as hearts at all, if biological categories are to be individuated functionally and functions are defined etilogically. Far from explaining the possibility of malfunction, the etiological theory can't even count a defective heart as a heart.<sup>15</sup>

Now, there are at least two ways for the etiologist to respond to this argument.<sup>16</sup> One way would be to retain the etiological account of function while abstaining from individuating organs and traits in terms of common function. But even if another principle of individuation, such as homology, might be enough to circumvent Davies' argument, it is not clear how much of the etiological account that survives without the claim that normativity applies with reference to a lineage for which selection has conferred a function to its members. If this lineage is allowed to contain homologous traits with very different functions (which would be all right if homology were to replace function in the individuation of biological traits), it is hard to see how the normativity that is to ground the possibility of malfunction will apply; no particular function will be identifiable as the normal one, since all kinds of performances may arise over time in a homologous lineage without loss of homology.<sup>17</sup> Another way to avoid Davies' conclusion is to question that a functionally individuated category must be so narrowly defined. Why not allow also nonfunctional hearts as members of the normatively constrained

---

<sup>15</sup> Davies 2000a, 26–30.

<sup>16</sup> A comment by Sören Häggqvist prompted the discussion in this paragraph.

<sup>17</sup> Compare the recent discussion in Neander 2002, especially p. 401.

group, as long as they belong to the selected lineage? Even though the original acquisition of function concerned only a narrow organ type, the descendants of organisms with that organ type would inherit the functional norm, also in case they had damaged organ tokens. The problem with this suggestion is that it appears to amount to little more than a stipulation to the effect that whatever is in the right successor relation to the selected type belongs to this type. But if the type is demarcated in terms of its success in competition with other, more or less similar types, there seems to be no reason to include malfunctioning tokens also in the selected success type and not just in the generic type that includes the success type as well as other less successful variants.<sup>18</sup> From the point of view of selected effect, it might even be said that the congenitally damaged heart is a heart in virtue of homology rather than function. The inclination to consider a damaged heart as malfunctioning rather than nonfunctional shows that a notion of biological normality is in play, but reference to selection does not explain this notion.

Davies' argument can be applied also to Buller's weak etiological theory, in which selection between variant traits is not required to have occurred. According to the weak theory, a heritable trait has a function if it was retained because of its contribution to fitness. A badly damaged heart unable to pump the blood will certainly not confer fitness, and it will thus not be a member of the class of items that in evolutionary history were maintained because of their contribution to fitness. So it is not a heart, and not malfunctioning either.<sup>19</sup>

There are also other objections to etiology, apart from the issue of normativity. One can question whether the etiological theory confuses the definition of function with an explanation of why there are functions. William Harvey's discovery of the heart's function is the standard example of a use of "function" which is independent of etiology. If Harvey was able to point out the function of the heart without any thought of its evolutionary origin, etiology cannot be part of the concept of function and has no place in its definition. To this Neander answers that we are concerned with a scientific term, and the

---

<sup>18</sup> Compare Davies 2000a, 29.

<sup>19</sup> Davies 2000a, 34.

meaning of such a term changes with changes in scientific theory.<sup>20</sup> To be sure, it is not uncommon that what is first proposed as an explanation of a phenomenon is later incorporated in its very definition. This is often the case for diseases. They are initially defined by their symptoms, but later on, when an explanatory hypothesis about the cause of a disease is sufficiently confirmed, the cause (for instance infection by a microorganism) may enter the definition of the disease. Similarly, on the etiological view, what was originally a hypothesis explaining why there are functions (namely an account in terms of natural selection) should subsequently enter the definition of function.

Such conceptual changes undoubtedly occur in the development of science. Whether this justifies the etiological theory's incorporation of an evolutionary explanation in the definition of function must be judged in relation to how the notion is used in biology. Considering the vast amount of cases where we have few clues as to the evolutionary origin of a trait, it seems overly restrictive to define function in terms of origin. An analysis that focuses on a trait's present use, which is what biologists usually determine before they attempt to investigate its evolutionary history, is clearly called for.

Another indication that the etiological analysis may not be entirely adequate is that it cannot account for functional traits the first time they appear. For instance, when a bacterium mutates in a new way which makes it resistant to some antibiotic, this trait will not have a function, according to the etiological theory, until it has been established in the population as a result of natural selection. A new mutation, says Neander, has no function: "at best it has accidentally beneficial effects" (1991b, 465). One may feel that even for the individual bacterium, already after a short time during which the new trait has protected its bearer from the antibiotic, this effect deserves to be called a function. If, on the other hand, in view of its accidental character, the trait is not considered to have a function, it is not quite clear why the mere fact that its accidental benefit will in subsequent generations lead to its becoming prevalent in the population should then give it the status of being a function. After all, one could view it

---

<sup>20</sup> Neander 1991a, 176. A similar point of view is found in Williams 1976, where some of the developments of the etiological theory in the 80s are anticipated.

as a case of the same accidental benefit, now spread among many bacteria. There has been selection for the trait, so it clearly is an adaptation. But it seems arbitrary to link the notion of function to that of adaptation; one would rather think that either the trait has a function when it has proven its value for its first bearer, or it never has one. The point is related to the objection that we are able to ascribe function without knowledge of evolutionary history. This is well expressed by Lowell Nissen, in a way that shows the closeness of these questions to the issue of normativity: “it would be surprising if, when current non-normative facts do not generate norms, past facts would; or if what an individual item does fails to generate norms, what many do, would.”<sup>21</sup> The point is not so much that a useful trait must be said to be functional from the very beginning – after all, intuitions may vary as to this – but rather that if one is certain that the trait initially has no function, why not hold that it doesn’t have one later on either? To stipulate that the concept of function can be cashed out in terms of adaptation is all right, but it is hard to see how this can preserve the normative aspect of function attributions. Adaptation is not usually taken to be a normative concept.

Less importantly, there has been much discussion about bizarre thought experiments. Neander (1991a, 179–180) imagines that some lions suddenly “pop into existence,” endowed with wing-like protuberances. Are these for flying? If they are not strong enough to enable flight, perhaps they have that function anyhow, but happen to malfunction? Or are they vestiges? Neander’s point is that without history, functional categories do not apply. (If the lions had a history, the questions would be answerable.) On the other hand, if we watch these lions killing and eating an animal, we won’t doubt about the function of their teeth. It is, unsurprisingly, hard to reach any particular conclusion from considerations of such scenarios.

These objections to the etiological account suggest that we face a choice: we may give up any distinctive notion of function, making do with adaptation, or we may look for an account that is independent of adaptation. The prevalence of function talk in biology seems to indicate that something other than adaptation is meant, so we had better try another account of function.

---

<sup>21</sup> Nissen 1993, 34.

## 2. CAUSAL ROLE FUNCTIONS

An account of biological function that, instead of linking the notion to considerations of evolutionary history, explicates it ahistorically in terms of the current performances of a system, was provided by Robert Cummins (1975). According to Cummins, the primary aim of functional ascriptions is not to explain a trait's presence or maintenance, but rather to explain its causal role, that is, what the trait contributes to some capacity of a system. What we do in functional analysis is to distinguish various systems that we wish to study, as for instance the digestive system or the circulatory system. Such a system has a complex capacity, which is to be analyzed into simpler capacities which together constitute the complex capacity of the system. That a component of the system has a function just means that it contributes to the complex capacity of the containing system. In the case of the heart, the containing system is the circulatory system, which has the capacity of transporting nutrients, oxygen, wastes, etc., and the heart's pumping of the blood contributes to this capacity. The circulatory system is the background relative to which the pumping of the blood is the heart's function.<sup>22</sup> An item's having a function amounts to its playing a causal role in regard to what the system does.

This theory makes it possible to ascribe functions to new properties, as in the case of a bacterium with a mutation protecting it from an antibiotic. It has the general advantage of not making functional ascriptions dependent on knowledge of evolutionary history. William Harvey's discovery of the heart's function is therefore less problematic,<sup>23</sup> and there is no need to deny functions to traits that do not have the selection history required by the etiological account.

---

<sup>22</sup> Because this background is the one usually presupposed, it often goes without mention. See Cummins 1975, 762.

<sup>23</sup> But not entirely unproblematic, as Wouters has shown. Harvey's question was not what role the heart's beating plays with respect to the circulatory system (though an answer to this question emerged as a result of his investigations), but rather why the heart has the structure it has (Wouters 1999, 103).



Amundson and Lauder (1994) have stressed the importance of causal role functions for such branches of biology as anatomy and morphology, where organism design is studied without regard to the evolutionary origin of traits. In view of what actually goes on in biological practice, their conclusion is that this notion of function cannot be eliminated “without also eliminating many key research questions.”<sup>24</sup> They opt for a pluralistic approach, in which etiological functions and causal role functions are both legitimate, in their respective contexts.<sup>25</sup>

It is possible, however, to go beyond pluralism and try to reach a unification of the concept of biological function by entirely eliminating etiological functions. This is attempted in Davies 2000b. According to Davies, the category of etiological function is not distinctive. Cases belonging to this category are in fact instances of causal role functions, so there is no reason to recognize etiological function as a separate category. In order to show this, Davies analyzes change brought about by natural selection. This can be conceptualized in terms of systemic change in line with Cummins’ theory.<sup>26</sup> The analysis proceeds in three steps, starting on the level of the population. The capacity of the population to change is analyzed into component capacities of its component groups, as for instance in a case where the organisms in one group have a trait that confers on them more reproductive success than what members of another group lacking the trait have. The capacities of these groups constitute the capacity of the entire system (namely the population’s capacity to change), and they are therefore functions relative to this background. As the change in the population may be brought about simply by differences in the proportions of the groups to each other (one group outreproducing the other), the groups may not be ascribed fully-fledged systematic functions (which would require interaction among the components), but even so they have causal roles and thus have functions of a kind. In a second step, focusing on the groups of organisms, one can investigate what explains their respective degree

---

<sup>24</sup> Amundson and Lauder 1994, 465.

<sup>25</sup> Melander 1997 also argues for the need for two different notions of function, one etiological and one in terms of causal role (though somewhat differently defined than in Cummins’ version).

<sup>26</sup> Davies 2000b, 90–93.

of reproductive success. Here, reproductive success is taken as the complex capacity to be analyzed. In a third step, we explain the capacities of the components found in the second step by functional analysis, identifying activities of simpler components in the organism's body that contribute to these capacities of the organism.

By this procedure, nothing but causal role functions are deployed, and Davies' claim is that every function ascription we might want to make from the standpoint of the etiological theory can be made also from within the causal role theory. There is then no reason for keeping a separate theory of etiological function, given that the causal role account also takes care of cases in which functions can't be ascribed on the basis of etiology (such as ahistorical function attributions).

The first step in Davies' account is a bit problematic in that it ascribes functions to groups of individuals in a population. In biological contexts, it is usually the individual organism that is taken to have functional parts (which in their turn may have functional parts). The ascription of functions to groups appears to imply a commitment to the view that species are individuals, and perhaps to rather controversial assumptions of group selection or species selection, since these views allow treating an individual organism or a group as part of a larger unit, itself an individual, and would therefore seem to legitimate functional or teleological attributions to groups as parts of such units. But Davies (2000b, 91) claims that his analysis merely presupposes that populations are viewed as systems, something biologists standardly do, and that this is independent of the further question of whether species should be taken as individuals. Neither group selection nor species selection is discussed by him at all.

That Davies sees nothing problematic in this can be explained by an important feature of causal role functions: their non-teleological nature. One does not have to hesitate whether something is appropriate for functional ascription, as long as it has a capacity that is identified as contributing to an encompassing system's capacity. There is no restriction to the effect that the system should be an organism or in any way biological. The system must have some complexity for functional analysis to be informative, but complexity can be found outside the domain of the biological. For instance, Davies (2000b, 94) mentions that the capacity of salt to dissolve in

water can be explained by appeal to the bonding capacities of the relevant kinds of molecules. Such molecular capacities satisfy the conditions of the causal role account for being functions. From this point of view there is thus no worry about whether an item normally would be said to be for the sake of something. A function ascription does not necessarily correspond to a pre-theoretical teleological attribution. The causal role theory is fully naturalistic, and it doesn't distinguish biological functions from any other functions. These aspects of the theory may seem attractive, but they also lead to some difficulties.

Before going into that, it should be noticed that the elimination of etiological functions proposed by Davies is not dependent on the first step of his procedure. The consideration above of the weaknesses of the etiological account already indicates the need for an analysis of function that is not contingent on selection history. Selection operates on traits that have advantageous functions. If we prefer to avoid "function" here, and use "effect" instead, it is hard to see why we should introduce "function" at a later stage, after the occurrence of selection. A trait whose effect has been selected for is an adaptation, and this effect could well be called a function, but we might as well say that a function preserved in selection is an adaptation. Furthermore, as Buller (1998, 511) has pointed out, the notion of "selection for" used for defining adaptation implies that the trait selected for contributes to the fitness of the organism, which amounts to its having a causal role function.<sup>27</sup> This argument pertains to the second and third steps of Davies' scheme, so the elimination of etiological function goes through even if we restrict functional ascriptions to the level of the organism and its parts, avoiding controversial issues about the individuality of species and selection on higher levels.

The eliminability of the etiological notion of function is less straightforward with regard to the approach of Peter Melander (1997). Melander, like Amundson and Lauder, claims that there are two distinct notions of function used in current biology. One is the etiological (defined as adaptation), the other one is linked to the

---

<sup>27</sup> Even though he shows that etiological function presupposes causal role function, Buller retains etiology in his account, presumably because he thinks that only selection history has the capacity to confer normativity.

concept of adaptiveness.<sup>28</sup> A trait is adaptive (relative to an organism in an environment E) just in case the trait “causally promotes the survival and reproduction of the organism in E.”<sup>29</sup> Adaptiveness is thus an ahistorical concept, indicating usefulness, whereas adaptation involves a causal history to the effect that a trait has been selected for and because of that is currently present.<sup>30</sup> The concept of adaptation is dependent on that of adaptiveness since a trait must have been adaptive at some point in history for it to have been selected for (thereby becoming an adaptation).<sup>31</sup> It might now appear that the etiological notion of function in terms of adaptation could immediately be eliminated, since for each case of applying the etiological notion to a trait, this trait must already have had a function as defined in terms of adaptiveness. But as Melander (1997, 87) points out, a trait can be an adaptation without (now) being adaptive, and this precludes the direct elimination of the etiological notion. For example, wings of birds are adaptations for flying, but if some birds are blown away to an isolated island where flight has no utility (there is plenty of food on the ground) and even is dangerous (a flying bird risks being blown out to sea), flying is not adaptive.<sup>32</sup> One is inclined here to question the intuition guiding this sort of example. If a trait no longer is adaptive, perhaps it no longer has a function. There might be some indeterminacy in the application of the notion of function in unusual circumstances. If on the other hand we ascribe the function of flying to the wings of the birds on the island, as Melander does, I

---

<sup>28</sup> Melander 1997, 90.

<sup>29</sup> Melander 1997, 82.

<sup>30</sup> The distinction between adaptation and adaptiveness is standardly found in recent discussions, for instance in Sober 1993, 84. Wouters (1999, 151–163) offers an excellent survey of different concepts of adaptation. Not so long ago, the concepts were not always rigorously distinguished, and terminology varied. For instance, Ruse (1973, 187) defined “adaptation” as what helps in survival and reproduction, which is how “adaptiveness” is now defined; cf. Baublys 1975 on Ruse’s assimilation of different concepts of adaptation. As Paul Needham has pointed out to me, when Ruse links the concept of function to that of “adaptation,” it amounts to much the same as Melander’s notion of function as adaptiveness. See also the discussion of function and benefit in Cohen 1978, 252–259.

<sup>31</sup> See Melander 1997, 87.

<sup>32</sup> Melander 1997, 72 and 87.

would think that we do this not primarily with reference to the evolutionary history of the trait, but rather because it is so obvious that flying is a capacity that wings make possible, and we take it for granted that flying normally is adaptive.

There is a presupposition of normality in play here, which may lead us to retain function attributions to traits no longer functional from a strictly adaptive point of view. It seems that such a standard of normality is what the etiological theory would need in order to account for normativity with regard to malfunction.<sup>33</sup> We shall return to this theme after looking at some standard objections to the causal role account, and how it can be restricted to meet such objections. Let it just be noted here that the only thing that prevents the reduction of Melander's etiological functions (adaptations) to his adaptiveness functions is that the former, though once adaptive, need not remain adaptive. If an explanation for the retention of function can be provided by examining standards of normality, this will perhaps not only make the etiological account of function superfluous, but also have a bearing on the problems pertaining to the causal role approach to function.

### 3. RESTRICTING CAUSAL ROLE FUNCTIONS TO BIOLOGICAL SYSTEMS

The most frequent objection to the causal role account is that it is too liberal in sanctioning functional attributions. It is therefore unable to differentiate biological systems from other complex systems, and as a consequence it can't explain why function attributions are plausible in biology but out of place in many other contexts. Neander (1991a, 181) claims that plate movements culminating in earthquakes would qualify for having causal role functions. Against this, Amundson and Lauder point out that the example does not satisfy Cummins' conditions for the use of functional analysis, since there is not enough complexity: the earthquake is really the same thing as the plate

---

<sup>33</sup> This is illustrated by the use of "typically" and "normal" in Sandra Mitchell's (1995, 45) defence of the etiological account's ability to accommodate normativity.

movements. “The geological structures which result in earthquakes might be complex, but the ‘organization’ of these structures *vis a vis* their explanation of the capacity of the earth to quake is not.”<sup>34</sup> Suppose, however, that there are non-biological natural systems complex enough for functional analysis to be worthwhile (disregarding the problem of how exactly complexity and organization should be characterized). It seems plausible to expect that such systems exist. The components reached in functional analysis of a system of that kind would then have functions. This much seems unproblematic. But contrary to what is the case with respect to biological systems (as well as for man-made artefacts, which I leave out of the discussion), there is no inclination to take these function ascriptions in any teleological or normative sense. We attribute a causal role function to an item in the non-biological system when it contributes to what the system does, but if the effect of the item changes in such a way that the system’s performance is altered, it is not a case of malfunctioning, except in an entirely and unproblematically metaphorical sense.

Adherents of the etiological approach consider the lack of normativity in causal role functions to confirm the advantage of their account. But in the light of the discussion above, it is doubtful whether etiology explains normativity more successfully. Both theories, in so far as they wish to find a place for the category of malfunction, presuppose a normative background against which it is distinguished from lack of function. The problem is less pressing for the causal role account, since from its point of view it is not a requirement that a functional ascription should explain why an item is there, but merely what it does.<sup>35</sup> Still, the problem seems to crop up here too, now in the form of a discrepancy between causal role functions within and outside of biology which is unaccounted for by the theory.

There are also biological counter-examples to the causal role account. Many organisms have the complex capacity to die of cancer. The outcome of the pathological process is a result of activities on the

---

<sup>34</sup> Amundson and Lauder 1994, 452; see also Wouters (1999, 121) for criticism of other putative counterexamples.

<sup>35</sup> See Cohen (1978, 252–258) for a discussion of the difference between functional ascriptions and functional explanations.

part of the tumour. This seems to satisfy Cummins' criteria for function attribution, but tumours, in Neander's words, "simply don't have proper functions."<sup>36</sup> Developing this example a little, we could imagine tumours to be some kind of parasitic organisms, and attribute functions to their activities. In this scenario, the ascription of functions to the tumour's activities would appear natural, whereas in the actual case, as tumours aren't organisms, their activities lack function. Functional attribution is thus tied to parts or processes of individual organisms.<sup>37</sup> As part of an organism, a tumour has no function, which suggests that for an item to have a function, it should not only contribute to the system's capacity, but in addition this capacity should be advantageous for the system, in contradistinction to the capacity to die of cancer.

A related objection to the causal role account concerns the distinction between functions and accidental effects. The sound of the heart's beating is a mere effect, its pumping the blood is not only an effect, but also a function. Now, a system in which the sound plays a role for what the system does is easy to imagine. In the context of a medical examination, the heart's sound can cause the doctor to begin a treatment. Functional analysis of this system reveals that the function of the heart is to produce sound. But we rather wish to say that the function of the heart is to pump the blood. The problem is not the perfectly plausible assumption that an organ can have more than one function (though this might be a bit troubling for such versions of the etiological account which stress that *the* function of a trait is *the* performance it was selected for); it is rather that it always appears to be possible to arbitrarily invent a function for anything, by imagining an appropriate system.

One way for the causal role approach to tackle the function/accident distinction is to distinguish the biological system from a wider system which includes social interactions and medical technology.<sup>38</sup> Biology is then separated as a special domain, relative to

---

<sup>36</sup> Neander 1991a, 181; see also Melander 1997, 54.

<sup>37</sup> Which suggests that if species, as it is sometimes claimed, are individuals (in more than a formal respect), it ought to make sense to attribute functions to their parts (that is, to the individual organisms as parts of their species).

<sup>38</sup> Maund 2000, 173.

which the heart's beating sound lacks function, although it may have a function in relation to the wider system.

What these objections show is that the causal role account must treat biology in a distinctive way if it is to preserve the peculiarity of the functional ascriptions in this domain, namely their teleological sense. Perhaps the apparent surplus of normative or teleological content in the notion of biological function should be treated as merely metaphorical. One way to resist that conclusion is to stipulate a further condition constraining what counts as an encompassing system. A suitable condition for this purpose is that such a system can be ascribed fitness (in the biological sense); in other words, that it is an organism.<sup>39</sup> As "fitness" is a fundamental concept in evolutionary theory, it is deemed unproblematic from a naturalistic point of view, and so a kind of natural good<sup>40</sup> is introduced that accounts for the teleological surplus distinguishing biological functions from other causal role functions in non-biological systems.

There are various versions of such restricted causal role theories. The most influential one in recent discussion is the propensity theory proposed by John Bigelow and Robert Pargetter. According to them, "something has a (biological) function just when it confers a survival-enhancing propensity on a creature that possesses it."<sup>41</sup> Functions, as survival enhancers, are thus closely related to fitness, which Bigelow and Pargetter consider to be an organism's disposition (based on its properties) to survive and reproduce, relative to an environment. On this view, when an organism accidentally fails to have the expected number of offspring (if for instance it has the bad luck to be struck by

---

<sup>39</sup> For an alternative way of restricting the domain for causal role functions, framed in terms of complex self-reproducing systems, see Schlosser 1998.

<sup>40</sup> This could also be called a goal, though not quite in the spirit of the early "cybernetic" approach to teleology as goal-directedness. This approach construed goals as preferred states, preserved under differing circumstances by feedback mechanisms. There is a vast literature criticizing the cybernetical view. Examples of critical objections: Schlosser (1998, 309) points out that some functional traits are not regulated by feedback; Wachbroit (1994, 584) argues that the cybernetic account can't distinguish between functioning and malfunctioning, because it only looks to the stable preservation of the end-state, a condition satisfied also when this end-state is worthless for the organism.

<sup>41</sup> Bigelow and Pargetter 1987, 192.



lightning), we do not for that reason revise our estimation of its fitness. An alternative view on fitness that characterizes it in terms of actual reproduction (leaving no place for bad luck) provides a way of retrospectively measuring fitness, but has the disadvantage of making the concept predictively empty.

Restating the propensity theory in the terms of the causal role theory, we get, roughly, the result that a trait's function is what it does that contributes to the organism's fitness (considered as a complex capacity).<sup>42</sup> Bigelow and Pargetter take function attribution to be forward-looking, in contrast to the backward-looking etiological view. As the organism's expected fitness is relative to the environment, we will have different functions for the same trait in different environments. Mitchell (1993, 259) points out that one consequence of this is that for an animal that is placed in a new environment (for instance a zoo) that differs from its natural one, we will have to redescribe the functions of some of its traits. By introducing new factors affecting the animal's survival chances, we could alter what functions the performances of its various organs have. To avoid such arbitrary variation, Bigelow and Pargetter add the condition that the environment relative to which a trait can be said to have a function is its bearer's natural habitat.<sup>43</sup> But, as Mitchell remarks (1993, 259), if this is taken to be the environment of evolutionary history, we are brought back to the etiological account. But rather than thus assimilating the propensity view to the standard etiological theory, Bigelow's and Pargetter's account shows a convergence to the modern history etiology of Godfrey-Smith (1994), in that the evolutionary history appealed to in determining the natural habitat is a history approaching present time, and doesn't necessarily go back to when the trait originally was selected for.<sup>44</sup>

---

<sup>42</sup> This is not quite right, since whatever an item does contributes to the fitness of its bearer, if only by lowering it. The contribution given by a functional item is rather to the capacity of having high fitness. But also this formulation is problematic, as will be seen below.

<sup>43</sup> Another way to handle this variation is to characterize function as a relational property, defined with respect to a selective regime. See Walsh 1996, 564.

<sup>44</sup> Compare Walsh 1996, 563.

It seems that at least some unification of the competing approaches has been reached. Davies' and Buller's arguments show that we can treat evolutionarily selected functions as causal role functions, and to avoid counter-examples based on the inappropriateness of function attributions in non-biological systems, the account is restricted to systems that survive and reproduce. In this way we reach a position with some resemblance to Bigelow's and Pargetter's, but also to the modern history version of etiology. Even though the details of these connections are not entirely clear, we seem to converge on a view according to which a function is something an item does in a system that enhances the system's capacity to survive and reproduce, given that there is some stability in history for this outcome.<sup>45</sup> This last condition (stability in history) would serve to rule out utterly accidental cases where some item, perhaps only once, happened to have a beneficial effect. It is not the etiologist's stronger demand for selection history.

However, all problems are not solved. As Wouters (1999, 129–130) remarks, there is a difference between attribution of function and of survival value. A trait has a function with respect to the organism, whereas survival value is more directly dependent on the environment. Survival value is a comparative notion, bringing the other members of the population into consideration. This suggests that the need for a function to be advantageous, illustrated by the cancer example, is not so easily met by the theoretical concept of fitness. Also, remember the birds blown to an island where flying is not advantageous. The wings are still for flying (we might wish to say), but flying is not adaptive; in fact, the individuals that use their wings to fly lower their chances to survive. Our functional ascription has something to do with normality or normativity, an issue that is still unresolved. In the warm environment of a zoo, the polar bears' furs retain the function of protecting them from cold, even though those with thin, badly functioning furs tend to survive longer. The claim that the furs are supposed to protect from cold depends upon what normally would be good for polar bears. We need to look somewhat closer at this notion of normality, which is the background

---

<sup>45</sup> In view of this convergence, recent accounts of function are perhaps less divergent than Ratcliffe 2000 suggests.

relative to which we can continue to ascribe functions to traits even when they cease to be advantageous.

#### 4. THE NORMAL, THE GOOD, AND ARISTOTELIAN FUNCTION

In an analysis of the biological notion of normality, Robert Wachbroit claims that it is not, as might be expected, the same notion as that of statistical normality. When it is said that hearts normally pump the blood, this is not a statement about the statistically average heart. “Suppose a calamity occurred in which most people’s hearts failed to circulate blood so that they needed an implanted medical device for this purpose. This would hardly undermine the statement about the heart’s function.”<sup>46</sup> Neither is biological normality an evaluative notion in an ethical sense. Wachbroit considers it to be a specifically biological notion.<sup>47</sup> He points out that it is not just a case of idealization, as for instance the use of point masses in physics is. A point mass has a position in space, but no spatial extension. It is an idealization, an abstraction made for theoretical purposes, which cannot actually exist. The case is different with biological normality: there are actually existing normal hearts.<sup>48</sup> Against the suggestion that normality can be captured in terms of function (so that the heart’s function defines what should be taken as its normal performance), Wachbroit argues that such a definition would be redundant, since biological function can’t be explicated without using the notion of normality (if only tacitly). Accordingly, the etiological theory needs

---

<sup>46</sup> Wachbroit 1994, 580. Millikan (1989a, 285; 1989b, 300) also stresses the importance of a non-statistical normality, though she accounts for it in terms of etiology.

<sup>47</sup> The separation of norm from average in Canguilhem’s influential investigation of the physiological concept of normality points in the same direction, though his analyses are quite complex and difficult to pin down. See Canguilhem 1978, especially p. 103. Amundson 2000 offers an excellent critique of the appeal to normality in biology and medicine, but he seems to mistake Wachbroit’s approach for the statistical view on normality.

<sup>48</sup> Wachbroit 1994, 588.

to demarcate the adaptations that are functions from those that are not. These latter are traits that are negative for the organism, but nevertheless can be given an evolutionary explanation.

In a similar vein, Richard Manning remarks that the attempts to naturalize function presuppose fitness as a goal, and that this is "an idea *exploited* rather than *explained* by the analysis."<sup>49</sup> He discusses the case of segregator distorter (SD) genes. By destroying their "rivals," they are able to get reproduced in a higher proportion of the gametes than should be expected by the laws of meiosis. The survival of SD genes is explained by their disrupting the normal outcome of meiosis, and so it seems to be a case of adaptation. Still we don't say that the function of SD genes is to disrupt meiosis, especially as they sometimes badly affect the organisms that bear them.<sup>50</sup> Now, an obvious move here is to assign to SD genes (or perhaps to parts of them) the function of disrupting meiosis for the sake of the SD genes themselves, and not for the organism. But as Manning points out (1997, 78), the fact that we have to search for the system which benefits from the trait shows that the etiological theory of function ascription does not accept as functions just any selected effects, but presuppose that they are of value. So the etiological theory should explicitly acknowledge that functions must contribute in a beneficial way to a biological system,<sup>51</sup> even if such a move threatens its naturalism.

The upshot of Wachbroit's and Manning's discussions is that naturalized accounts of function depend on an unanalyzed conception of normality or goal, which is not explained by statistically examining biological performance or reconstructing selection history. Against this, it might be argued that assigning pumping as the heart's normal function in Wachbroit's imagined scenario is still a case of statistical normality, despite the massive majority of malfunctioning hearts. The statistically relevant group can be taken to be composed not only of people after the occurrence of the calamity, but also of all people having existed before. A better example of the difference between statistical and biological normality would be the acorn. Its

---

<sup>49</sup> Manning 1997, 71. See also Machamer 1977 for an analysis that drives a wedge between a selection process and the benefits it leads to.

<sup>50</sup> Manning 1997, 75; Godfrey-Smith 1994, 348.

<sup>51</sup> As is done in Godfrey-Smith 1994, 349.

biologically normal tendency is to become an oak, but normally (statistically speaking) acorns rot on the ground.<sup>52</sup> Gerhard Schurz (2001) attempts to counter such examples and maintain a connection between a trait's biologically normal function and its statistically normal performance. His strategy, applied to acorns, is to consider the individual acorn's normal function to consist in having the capacity to become an oak, rather than in actually becoming an oak. Whereas becoming an oak is rare for acorns, having the capacity is common. Therefore the average acorn really does perform what it is supposed to do, even though it doesn't grow to a tree. Its capacity is still there. Schurz' appeal to the capacity as normal function saves the connection to statistical normality, but it comes with a cost: it seems to abolish the possibility of malfunction. If an acorn rots, it still functions. Under what circumstances does it malfunction? Only of an acorn that doesn't have a capacity to grow could it be said that it malfunctions. But even in such a case, who knows what could be done to get it to germinate, perhaps with the aid of advanced biotechnology? As long as we are defining malfunction with respect to a capacity, it is hard to exclude possible circumstances relative to which even the most damaged acorn will have the capacity to grow. It is only by stipulating what counts as *normal* circumstances that malfunction is possible, and this normality does not seem to be of the statistical kind. For even though most acorns deemed to be unable to germinate are not brought to the biotechnologists, it is not this statistical fact that counts, but merely their capacity to grow given an appropriate (and perhaps ideal) environment, and this capacity is the same for all non-germinating acorns, whether there are any biotechnologists around or not.

To sum up, the biologically normal is thus what happens often enough (but not necessarily in the majority of cases) and is considered good for the organism. Determining the normal course of things is therefore dependent upon having this good in view. And, as was said in the previous section, what is taken as good is not necessarily fit in the technical sense of evolutionary theory, though it

---

<sup>52</sup> Compare Millikan 1989a, 285. Bergmann (1968, 222) uses this example for somewhat different purposes.

typically (normally) is. Inserting “typically” here mirrors the flexible retention of function attribution in cases where no advantage comes from exercising the function. The notion of function in biology is closely related to considerations of normality and benefit, and the attempts to account for it independently of such considerations do not appear to be successful. But this amounts to saying that biological function is a teleological notion that does not yield to naturalistic reduction, since at least the common kind of naturalism does not admit of non-reduced evaluative notions.<sup>53</sup>

In this criticism of naturalized function theories, a similarity between the etiological and the causal role approaches comes to view. The causal role analysis on its own does not demarcate any special domain in which functional ascriptions are to be employed, and critics generally consider this to show that its explication of “function” is excessively liberal. An additional constraint on the theory is needed to save the intuition that functions are more appropriately attributed to living systems and artefacts than to other systems. The etiological account is usually taken to fare better in this respect: if functions are selected effects, they are kept in the right domains, namely the biological (natural selection) and the artefactual (intentional selection). However, it turns out that even on this view a system for which the function serves has to be presupposed, as all selected effects aren’t functions.

At this point it is interesting to take a look at the Aristotelian tradition in the philosophical analysis of function, nowadays a minority position represented among others by Richard Sorabji (1964), Andrew Woodfield (1976), Mark Bedau (1991), and Barry Maund (2000).<sup>54</sup> I will not try to estimate to what extent the accounts of these authors correspond to views actually held by Aristotle. An Aristotelian view of function is here taken to be one that places contribution to something good for the organism in focus, without

---

<sup>53</sup> A theory of biological function from the point of view of a naturalism admitting natural norms is offered by Christensen and Bickhard 2002 (see especially p. 14).

<sup>54</sup> See also McLaughlin 2001, 14: “Even the naturalist needs a bit of Aristotle to reconstruct our functional attributions.” McLaughlin’s book contains penetrating analyses relevant to all problems under discussion in the present essay.

trying to reduce this good to something else. The above-mentioned authors differ in various respects, but they have in common the view that a functional statement implies that there is something good for the bearer of the functional trait to which the trait contributes. In Woodfield's analysis (1976, 208), biological functions should be understood in terms of an item X (part of a system S) doing its characteristic activity A because A characteristically and normally contributes to F, which is (normally) good for S (where F is an activity of X causally generated by A).<sup>55</sup> On this theory, "it is natural to describe the beings served by functions as 'beneficiaries'."<sup>56</sup>

Bedau likewise connects function to value, taken as that which is in the interest of the beneficiary. He points out that the value involved in ascriptions of biological functions is not generated by natural selection per se, as there are cases of natural selection which are not considered to produce functions. For example, "populations" of clay crystals have four features characteristic of natural selection: crystals reproduce (big crystals break and the pieces become "seeds" for new crystals), they show enormous variation ("flaws"), variant traits are inherited by the "offspring," and the inherited traits can affect the rate at which crystals of that kind proliferate in the "population."<sup>57</sup> Distinctive of such cases of natural selection in which no functions emerge is that they lack beneficiaries, there being nothing for which a certain outcome is good. On Bedau's view, only something living can be a beneficiary, that is, have interests.<sup>58</sup> If it is the case that some crystallization processes are instances of natural selection, this reinforces the point made by Wachbroit and Manning that without additional conditions, the etiological account is no more able to demarcate an appropriate domain of functional ascription than the causal role account is.

---

<sup>55</sup> In light of what has been said above it is interesting to note the frequent appeal to the "normal" and "characteristic" in this explication.

<sup>56</sup> Woodfield 1976, 122.

<sup>57</sup> Bedau 1991, 651–653.

<sup>58</sup> Bedau recognizes that there are borderline cases, such as viruses, of which it is hard to say whether teleology applies to them (Bedau 1992, 792). He also makes clear that having an interest by no means implies having any awareness of the interest.

*Biological Functions in a Kantian Perspective*

Maund (2000, 168) offers the following Aristotelian account of biological function:

The function of X is to do Y if and only if a) X is the kind of thing that results in Y, or enhances the capacity to do Y; b) Y is the mark of successful performance for [organism] O, i.e. the mark of what O is specifically good at, is successful at; or, Y is the necessary condition for something else's occurring which is the mark of success for O; or, ...and so on.

The idea is that the various kinds of organisms are good at their typical “tasks” (not only surviving and reproducing, but also more specific activities). A part of an organism (X in Maund’s scheme) has a function by contributing to the organism’s performance. Maund’s account is certainly not unobjectionable; for instance, since there are sometimes back-up systems, it might be better to weaken condition b), so that Y’s contribution is something less than a necessary condition. But details aside, what is clear is that no attempt is made to evade a commitment to values and beneficiaries. This is the reason why this kind of analysis is often deemed inappropriate from a naturalistic point of view. In Godfrey-Smith’s words (which are somewhat in tension with his own account’s appeal to the beneficial contributions of traits), “an appeal to goals is certainly a step backwards.”<sup>59</sup> It is interesting to note that, according to James Lennox, recent studies of Aristotle’s teleology “stress its naturalism.”<sup>60</sup> In Lennox view, Aristotle tried to find an intermediate position between the reductive materialism of some presocratic thinkers and Plato’s theocentric teleology. Analogously, to decide whether modern accounts of function inspired by Aristotle’s teleology are to be called naturalistic, one must assess whether contemporary naturalism is overly reductive, and whether other forms of naturalism are possible. In any event, if there is any plausibility in the conclusions reached above as to the shortcomings of the etiological and the causal role approaches, some appeal to what is good for the organism seems to be needed.

---

<sup>59</sup> Godfrey-Smith 1994, 348; see also Manning 1997, 77.

<sup>60</sup> Lennox 2001, 259.



### *Biological Functions in a Kantian Perspective*

In the next sections I will construe Kant's teleological perspective on the organism as to a large extent compatible with the Aristotelian approach to functions. The distinctive feature of Kant's view is that it combines the Aristotelian approach with a reflective stance, in which teleology is seen as an *epistemic presupposition* constitutive for the study of life, rather than a definite *ontological* commitment. In this way biology is delimited as a form of inquiry based on a supposition of teleology. The issue of naturalism is bypassed, in so far that a teleological perspective constitutes biological science and the possibility of grasping the objects it studies, while on an ontological or philosophical level teleology is neither affirmed nor denied. The Kantian point, on this reading, is that, regardless of whether a reductive naturalism (that doesn't allow values ascribed to organisms) is ultimately true or false (and regardless of whether this is possible for us to determine), without a teleological perspective biological science is not possible.<sup>61</sup>

## 5. KANT'S NOTION OF THE ORGANISM

The Kantian perspective on function as presented here draws on key aspects of Kant's actual position, but does not require the validity of all his views on biology or philosophy generally. A balance will be sought between the demands of exegesis and the aim of connecting Kant to contemporary philosophy of biology. First it may be useful to give a short and simplified sketch of how the question of biological explanation emerges in the context of Kant's critical philosophy. In the *Critique of Pure Reason*, Kant had delineated a system of categories and principles flowing from them which together put a

---

<sup>61</sup> There are points of contact between Kant's and Daniel Dennett's views on teleology, as shown for instance by Dennett's claim (1993, 224) that "biology depends [...] on adopting the intentional stance towards the evolutionary process itself." But see Ratcliffe 2001 on the tension between naturalism and Kantian perspectivism in Dennett's account. Kant's view also has some remote resemblance to the claim in Rosenberg 1994 that biology, in contradistinction to physics, is an instrumentalistic science, though Rosenberg reaches this conclusion from entirely different considerations.

*priori* constraints on spatio-temporal experience, in so far as it is to be objective. Most importantly, the principle of causality, according to which every event has a cause, is a necessary condition for the possibility of objective experience, and on this basis the search in experience for empirical laws of nature is warranted. In the *Critique of the Power of Judgment*, originally published in 1790,<sup>62</sup> Kant turned his attention to teleological explanations in biology.<sup>63</sup> The problem he confronted here was that there are features of organisms that appear to be intractable for the kind of explanations in terms of causal laws appropriate for ordinary physical objects (what Kant calls “mechanism”). Organisms are characterized by their capacities for reproduction, growth (which involves the radical transformation of ingredients taken up from outside), and regeneration (for instance compensation for the loss of a part), capacities which in Kant’s view are utterly different from anything else met with in nature (CJ 371–372). What leads us to view objects with such capacities in teleological terms is that the causal relations they exhibit can be seen “as law-governed only if we regard the cause’s action as based on the idea of the effect, with this idea as the underlying condition under which the cause itself can produce that effect” (CJ 367).<sup>64</sup> It is thus the need for explanation in terms of law that, in the absence of mechanistic explanations, forces us to adopt a teleological view of organisms, in which a purpose (the “idea of the effect”) is taken into consideration. Even though organisms like all objects of experience are subject to the causal principle, this alone does not take us very far in explaining their characteristics. Seen from a purely causal point of view, there is no explanation (or “law”) for how matter comes together in the ways characteristic for organisms.

---

<sup>62</sup> Henceforth this work is referred to as CJ. Page numbers indicate the pagination of volume 5 of the *Akademie* edition of Kant’s works, also given in Kant 2000. When below I refer to the *First Introduction to the Critique of the Power of Judgment*, a text not published by Kant himself, page numbers refer to volume 20 of the *Akademie Ausgabe*. The *First Introduction* (referred to as FI) is included in Kant 2000.

<sup>63</sup> I use the terms “biology” and “organism” for convenience, though these terms are not used in CJ. The first term was almost never employed at that time in its modern sense, and for the second one Kant mostly has “*organisiertes Wesen*.”

<sup>64</sup> Pluhar’s translation, which is more perspicuous here (Kant 1987, 244).

*Biological Functions in a Kantian Perspective*

For if one adduces, e.g., the structure of a bird, the hollowness of its bones, the placement of its wings for movement and of its tail for steering, etc., one says that given the mere *nexus effectivus* in nature, without the help of a special kind of causality, namely that of ends (*nexus finalis*), this is all in the highest degree contingent: i.e., that nature, considered as a mere mechanism, could have formed itself in a thousand different ways without hitting precisely upon the unity in accordance with such a rule. (CJ 360)

The organism with its typical features (reproduction, growth, regeneration) exhibits a kind of causal structure where effective causes seem to be reciprocally conditioned by their own effects. The organism is thus in a way both cause and effect of itself. In order to handle this situation we resort to the *nexus finalis* known to us from the production of a work of art. So it might appear that Kant just presupposes that organisms are artefacts. But in presenting the criteria for calling a thing a “natural purpose” (*Naturzweck*, Kant’s term for a teleologically conceived organism) he explicitly adds the requirement that the thing produces itself. The notion of an external agency forming the thing according to a plan (which is how the production of artefacts is conceived) is replaced by an internal teleology in which each part “must be thought of as an organ that *produces* the other parts (consequently each produces the others reciprocally)” (CJ 374). For something to be a natural purpose it must be “an *organized* and *self-organizing* being” (CJ 374), and this self-organization is not a matter of replacing the external agent with an internal, conscious agency. Even though we describe the organism with locutions borrowed from the sphere of intentional production, “one says far too little about nature and its capacity in organized products if one calls this *an analogue of art*” (CJ 374); “strictly speaking, the organization of nature is therefore not analogous with any causality that we know” (CJ 375).

With this characterization of the organism as a self-organizing being that must be understood teleologically, we seem to stand firmly on Aristotelian ground. In spelling out the relevance of Kant’s views to contemporary discussions of biological function, two main questions have to be addressed. First, is there any important difference between Kant’s account and an Aristotelian analysis of function? And second, can anything of the pre-Darwinian Kantian

view remain viable, considering the huge advances in the life sciences in the last two centuries? The first question will be discussed in section 6, the second one in section 7.

## 6. NORMATIVE FUNCTION AS REGULATIVE MAXIM

No attempt will be made here to enter the complex issues in the interpretation of Aristotle's thought, so Kant's position will be compared to a rather schematic Aristotelian picture. In that picture, organisms are conceived of as self-organizing beings in possession of interests. There are typical goods from which they can benefit, ranging from species-universal values like survival and reproduction to the various species' particular values. These goods are taken to be objective features of nature, basic for biology, so one should not attempt a reduction of them. Attributions of function are to be explicated in terms of such goods (as seen for instance in Maund's analysis above).

One point where Kant differs from this picture concerns the locutions used in formulating functional attributions. Kant typically uses intentional locutions, like end or purpose (*Zweck*) and aim or intention (*Absicht*). This gives the impression that his teleology is based on the model of intentionally designed artefacts requiring an external agency. But as we have seen, he also stresses that natural purposes are natural, self-organizing beings. Much of Kant's discussion of biology in CJ centers around his claim that it is necessary to view organisms both as designed and as natural, i.e. non-designed. On the face of it, this seems to be an impossible combination. To reconstruct how Kant avoids a contradiction is a complex task, and I shall here only comment on the following two strands in his discussion. (i) Intentional locutions have a normative connotation; (ii) teleology has only regulative status. Point (i) is clearly expressed in the *First Introduction to the Critique of the Power of Judgment*, and it serves to bridge the gap between

descriptions of teleology in terms of intentions and in (Aristotelian) terms of what is good for the organism.<sup>65</sup>

A teleological judgment compares the concept of a product of nature as it is with one of what it *ought to be*. Here the judging of its possibility is grounded in a concept (of the end) that precedes it *a priori*. There is no difficulty in representing the possibility of products of art in such a way. But to think of a product of nature that there is something that it *ought to be* and then to judge whether it really is so already presupposes a principle that could not be drawn from experience (which teaches only what things are).<sup>66</sup>

Viewed in this way, attributing an end or a purpose to something is tantamount to judging it as subject to a norm or standard. This applies to a work of art as well as in biology. Intentional design is not just a matter of having a plan, but also of letting the plan serve as a standard for how the designed thing ought to be.<sup>67</sup> Similarly, a biological function is not just any effect of one of its parts that is to some advantage to the organism, but one that contributes to what it normally does, in the biological sense of normality which leans on a normative standard. This normative perspective is reminiscent of what Wittgenstein says in the *Philosophical Investigations* about the machine as symbol of its own mode of working.<sup>68</sup> When we think of a machine in this way, we consider its parts as if they could not move differently, as if they never could break down. We regard the machine in a timeless fashion, with its predetermined movements “already present” in advance.<sup>69</sup> This resembles what occurs when we ascribe functions against the background of biological normality. The function of *the heart* (timelessly and normatively conceived) is to pump the blood, and it is with this “symbol” (to use Wittgenstein’s word) that the actual performance of a real heart is compared when it

---

<sup>65</sup> Here I closely follow Hannah Ginsborg 2001, who stresses the significance of this passage and of a normative interpretation of Kant’s teleology.

<sup>66</sup> FI, Ak. 20:240.

<sup>67</sup> Compare Ginsborg 2001, 250.

<sup>68</sup> Wittgenstein 1977, §§ 193–194.

<sup>69</sup> Wittgenstein 1977, § 193.

is judged to function or malfunction. The important aspect in the present context is not that the machine has been intentionally built, but that its performance is measured against a norm. To exemplify, Kant points to our differing attitudes towards an eye and a stone: we “judge that it [the eye] *ought* to have been suitable for seeing,” and “this ought contains a necessity which is clearly distinguished from physical-mechanical necessity,” but we don’t regard a stone in that way, even if it is used for a purpose, for instance to build something. “I cannot on that account say that it ought to have served for building.”<sup>70</sup>

The normative interpretation of Kantian teleology brings it close to the Aristotelian conception of what is good for the organism. This latter notion can easily be taken as a norm: what the organism typically is good at is also what it ought to do. It is the standard by which we estimate whether its organs are functioning or malfunctioning.

As to point (ii) above, concerning the regulative status of teleology, this is what really distinguishes the Kantian account from the Aristotelian. According to Kant, when we judge an object teleologically (as for instance when we say what function an organism’s heart has), our statement does not ascribe an objective property to the object. It expresses a “way of judging,” a perspective necessary for the study of nature with respect to “a special class of its objects” (CJ 382). Kant deploys his distinction between the determining and the reflecting employment of the power of judgment. The former ascribes objective properties to objects (in accordance with the principles constitutive for empirical cognition that are derived from the categories), whereas the latter uses maxims for investigating nature where no laws in terms of objective properties are to be found. These latter maxims are said to be merely regulative, helping us to orient ourselves in the diversity of experience. Kant recognizes a necessity for regulative maxims in natural science generally, for instance principles of parsimony and simplicity (see for instance CJ 182). But the organism is a very special case: here regulative maxims are used for judging one thing internally. Such judging is called for because features such as reproduction, growth,

---

<sup>70</sup> FI, Ak. 20:240.

and regeneration, uniquely found in organisms,<sup>71</sup> show a regularity that has to be accounted for somehow. In the absence of mechanical explanations, normative rules “bring lawlike order to the otherwise incomprehensible diversity of the organic world.”<sup>72</sup>

In addition to its characteristically Kantian focus on the rules for cognizing objects rather than on the object as such, the distinctively unAristotelian move of relating the norms to us rather than to the object reflects an unwillingness to attribute norms or values to natural objects themselves. This attitude is shared by many modern philosophers and explains why the Aristotelian approach is a minority position in the contemporary debate on functions. For Kant, abstaining from ontological commitments also has the benefit of cancelling metaphysical speculation on the ultimate origin of living things. A teleological judging of nature can thus borrow intentional language (and even use the idea of divine design as a heuristic tool)<sup>73</sup> without theological implications. The other side of this coin is that Kant also considers attempts to explain the origin of organisms (the origin of life as well as the development of an organism) mechanistically as illegitimate metaphysical speculation:

For it is quite certain that we can never adequately come to know the organized beings and their internal possibility in accordance with merely mechanical principles of nature, let alone explain them; and indeed this is so certain that we can boldly say that it would be absurd for humans even to make such an attempt or to hope that there may yet arise a Newton who could make comprehensible even the generation of a blade of grass according to natural laws that no intention has ordered (CJ 400).

Kant is quick to add that it would be presumptuous to claim that organisms *could* not have a purely mechanistic origin, “for how could

---

<sup>71</sup> Bedau’s crystals might be proposed as a counter-example here too. But the task of providing chemical and mechanical explanation is presumably very much easier in this case than for the features of organisms. Interestingly, Kant discusses formation by crystallization in CJ 348–349. He concludes that such formations are to be explained mechanically, despite their beauty which gives the appearance of design.

<sup>72</sup> Ginsborg 2001, 253.

<sup>73</sup> As for instance in CJ 399–400.

we know that?” (CJ 400). Disregarding the problematical status of the impossibility claimed for any mechanistic explanation of life (a claim easy to assent to if limited to the mechanistic theories available in Kant’s time), the upshot in the present context is that in biology, as conceived by Kant, mechanism as well as teleology are valid only as heuristic maxims.<sup>74</sup>

Kant defines the teleological principle for judging organisms as follows: “An organized product of nature is that in which everything is an end and reciprocally a means as well” (CJ 376). Whereas modern definitions of function often refer to causal contributions to what a system does, Kant’s teleological perspective permits him to express such contributions in terms of means and ends. Advantage for the organism itself, not for something external to it, is here understood as constituting the end (internal purposiveness in contrast to relative purposiveness; see CJ 367–369). The modern discussion (except for some Aristotelian exceptions) is conditioned by the aim of reducing teleology to straightforwardly causal terms. I have argued (in sections 1–4) that teleology tends to crop up somewhere in the proposed analyses, at least implicitly. Though Kant rarely if ever uses the word “function” in the sense it has in contemporary philosophy of biology,<sup>75</sup> a Kantian definition of function can be proposed:

**KF**            The function of x in organism O is y iff x is a part of or a trait of O, x is a means to y, and y is one of O’s ends or a means to such an end.

Of course KF is of no use for the reductive purposes to which definitions of function are usually offered.<sup>76</sup>

To return to the comparison of Kant’s and Aristotle’s views, I would suggest that from Kant’s point of view, Aristotelian functions (as defined more or less in Maund’s fashion, or according to KF) are to be seen as valid at the *object level* of biological investigation. The denial of their objectivity only takes place at a further *meta level* reflection. For the biologist, organisms exhibit functions as a matter of course. It is in an ontological perspective, which is of no great

---

<sup>74</sup> I will return to the issue of the inexplicability of organisms in the next section.

<sup>75</sup> But see CJ 375n.

<sup>76</sup> Compare Zumbach (1984, 22), with discussion along similar lines.



importance for the practice of biology, that functions are assigned merely regulative status. The causal principle is accorded objective status in the ontological perspective,<sup>77</sup> but on its own it is of little use for the biologist, because without complementation from the teleological maxim there just are no biological objects to study. In this way, teleology, though regulative, can be said to be constitutive for biology, as a condition for the possibility of biological objects (organisms).

This sounds doubly paradoxical. How can a regulative principle be constitutive? And how could teleology, a mere way of judging, be said to provide objects? The first paradox is alleviated by a distinction of levels. On the ontological level, we have no ground to affirm the objectivity of norms. Only what is necessary for the possibility of experience in general is objective from this very austere perspective.<sup>78</sup> Biology is not necessary for the possibility of experience and biological normativity is therefore merely regulative. For biology to be possible, on the other hand, a teleological conception of the organism must be in place, so for this special science teleology is constitutive.<sup>79</sup>

---

<sup>77</sup> The ontological perspective corresponds to what Kant calls empirical realism. At a further level of reflection, Kant subordinates ontology to the transcendental conditions for the possibility of experience.

<sup>78</sup> For the details, see *Critique of Pure Reason*.

<sup>79</sup> Compare FI, Ak. 20:235–236: “But now we find among the products of nature special and very widely distributed genera, which contain within themselves a combination of efficient causes that we must ground in the concept of an end, even if we wish to employ only experience, i.e., observation in accordance with a principle suitable to their inner possibility. If we wished to judge their form and its possibility merely in accordance with mechanical laws, in which the idea of the effect must not be taken as the ground of the possibility of their cause, but vice versa, then it would be impossible to obtain even so much as an experiential concept of the specific form of these natural things which would put us in the position to move from their inner disposition as cause to the effect [...]: it is clear that if there are such products of nature, it would be impossible to even investigate their character and their cause in experience (let alone to explain them by reason) without representing their form and causality as determined in accordance with a principle of ends.” (Guyer’s and Matthews’ translation, slightly modified.)

The second paradox is eased in considering that the objects constituted by teleological judging are epistemic objects. This means that from the ontological perspective, nothing is added by teleology. All the physical and chemical processes taking place in a dog are (fortunately) still there when we stop doing biology. In principle they can be the object of study also without recourse to teleological considerations, though it seems that we would then have no reason to single out the dog as a significant unit of study. To be sure, even if not identified normatively by means of teleology, the dog as a chunk of matter could be individuated nonetheless, and its particular physical and chemical properties could be investigated. But without the normative functional context, the results of such an investigation, however interesting, would be very far from what we call “biology,” and the chunk of matter under study would be conceived of in a way entirely different from how we consider biological objects. In modern biology, where often purely physical and chemical studies are conducted, there is always in the background, explicitly or implicitly, a functional view of the biological entities concerned, which is what gives the enterprise its sense. This can be confirmed by a glance in any paper on molecular biology. Biochemical factors are consistently described with reference to their role in maintaining or disturbing the normal functioning of the organism.<sup>80</sup> Clark Zumbach (1984, 139) offers an example from biochemistry to illustrate Kant’s view:

[A] biochemist is free to view the binding of the amino acyl synthetase molecule to loop I of the tRNA molecule as but an event in an indefinite causal chain of events. [...] Yet when the biochemist views such an event as a part of an organism’s formative activities, the idea of design is a necessary universal for the reflective capacity of judgment in his or her attempts to

---

<sup>80</sup> The first sentence of the abstract of a randomly chosen research communication (Walker, Downes and Leslie 2001) illustrates this: “The PTEN (phosphatase and tensin homologue deleted on chromosome 10) tumour suppressor is a phosphatidylinositol 3,4,5-trisphosphate [PtdIns(3,4,5)P<sub>3</sub>] 3-phosphatase that plays a critical role in regulating many cellular processes by antagonizing the phosphoinositide 3-kinase signalling pathway.”

*Biological Functions in a Kantian Perspective*

understand these wholes. And this idea is not reducible to the mechanistic point of view.

Kant, writing before the era of molecular biology, exemplifies his view with a remark on the practice of anatomists:

It is well known that the anatomists of plants and animals, in order to investigate their structure and to understand for what reason and to what end they have been given such a disposition and combination of parts and precisely this internal form, assume as indispensably necessary the maxim that nothing in such a creature is *in vain* [...]. In fact, they could just as little dispense with this teleological principle as they could do without the universal physical principle [i.e. the causal principle], since, just as in the case of the abandonment of the latter there would remain no experience at all, so in the case of the abandonment of the former principle there would remain no guideline for the observation of a kind of natural thing that we have conceived of teleologically under the concept of a natural end. (CJ 376)

When comparing the roles of the teleological and the causal principles, it is seen that both serve as transcendental conditions, in other words constitutive principles, but on different levels: the former for anatomy, the latter for experience in general.<sup>81</sup> Since only the latter is absolutely indispensable (after all we could have objective experience without biological science), and also because of the problematic value ascription implicit in functional attributions, teleology's status is merely regulative. As indicated in the quote above, the locus for the application of the teleological principle is the kind of natural things that are "conceived of teleologically under the concept of a natural end [*Naturzwecke*]," that is, the organism. Hence the teleological concept of organism is constitutive for biology, and with it follows the methodological principle of searching for functions for every trait.

I have construed Kant's position with the aid of a distinction of levels suggested by, though not explicitly made in the *Critique of the*

---

<sup>81</sup> See also the discussion in McFarland 1970, 109–110.

*Power of Judgment*.<sup>82</sup> That Kant's view is somewhat more complex than depicted so far can be seen in his occasional remarks to the effect that the teleological principle is a regulative maxim not only from a meta level perspective, but also as viewed from the point of view of biological practice. This can be accommodated in the present scheme in the following way. What is constitutive for biology is the teleological concept of the organism. The heuristic maxim that every part of an organism has some function serves as a guideline for research and is in that sense regulative. It is a rule for viewing every feature discovered or examined as contributing to the general good of the organism. If we distinguish the philosophical meta level from the biological object level, we can put it like this. From the meta level perspective the teleological concept of organism is merely regulative (it pertains to the reflective power of judgment, not to the determinative, in Kant's terminology), but it is constitutive for the possibility of biology. From the object level perspective of biological practice, the concept of organism with the concomitant concept of biological functions has constitutive status (it delineates an objective field of study), and there are in addition heuristic principles of methodology which have regulative status in biological investigation. In this way claims about regulativity and constitutivity are seen to be legitimately made at different levels, and as long as the relevant level is kept in view confusion can be avoided.<sup>83</sup>

---

<sup>82</sup> One hint in this direction is in CJ § 58, where Kant says that he will show later in his book that plants and animals must be considered teleologically, "in accordance with the principle of realism" (CJ 349). In the subsequent discussion of teleology (CJ §§ 72–73), however, realism and idealism (the positions affirming and denying natural teleology respectively), are both rejected because of their common presupposition that teleology is an objective or determinative concept. Kant's divergent pronouncements can be reconciled with the aid of the distinction of levels here proposed. Teleology is treated realistically on the object level, though rejected (together with its contrary, idealism) on the meta level.

<sup>83</sup> This can be compared to some remarks of Martin Heidegger's concerning the presuppositions for zoological investigation. In the context of a discussion of von Uexküll's ethology, Heidegger describes the status of a certain proposition: "The animal is poor in world". This proposition "involves an antecedent determination of what belongs in general to the essence of the animal, that is, a determination of the field within which any positive

## 7. CONCLUSION: KANT AFTER DARWIN

Kant's reasons for rejecting the very idea of "a Newton who could make comprehensible even the generation of a blade of grass" (CJ 400), while at the same time holding that organisms may (or may not) ultimately have a mechanistic ground, are complex and difficult to pin down. In what sense are organisms mechanistically inexplicable? According to Peter McLaughlin's interpretation of Kant, it is a human peculiarity that causality is always treated as implying the explanatory reduction of wholes to parts, something that does not work with regard to organisms, as these show a holistic causality in which the parts would have to be explained by the whole. As we are unable to understand such an explanation, the concept of purpose serves as a regulative replacement.<sup>84</sup>

On Hannah Ginsborg's interpretation, the insufficiency of mechanism consists in the fact that mechanical laws (in a broad sense including chemical laws) cannot on their own account for the origin of an organism. To explain this we have to "appeal to [some] particular arrangement" of matter; we can't do it "solely by appeal to the laws governing the (unorganized) matter out of which it comes to be."<sup>85</sup>

Each of these interpretations has its problems, and I will not attempt to adjudicate between them here.<sup>86</sup> Regardless of how exactly to construe Kant's claim, the prospects for mechanistic explanation<sup>87</sup> of organisms are much better than in his time. Isn't Darwin the Newton of the blade of grass? I would suggest that it is rather in molecular biology that our Newton might appear. After all, Darwin's project was to explain how species could diversify (and progress in

---

investigation of animals must move. [...] The proposition does not derive from zoology, but it cannot be elucidated independently of zoology either. It requires a specific orientation toward zoology and biology in general" (Heidegger 1995, 186–187). The proposition discussed by Heidegger differs from Kant's principle of teleology, but the role attributed to it (determining the field of zoology) resembles the function of teleology.

<sup>84</sup> McLaughlin 1990, 152–179.

<sup>85</sup> Ginsborg 2001, 242.

<sup>86</sup> But see the next essay for more discussion.

<sup>87</sup> With this is henceforth meant non-teleological explanations in terms of physics and chemistry, not just in terms of mechanics in a narrow sense.

adaptation) out of one common root. An account of life in mechanistic terms would rather seem to be the business of chemistry and physics (though of course with the assistance of a study of natural selection). In view of the enormous gaps in our understanding of most generative processes,<sup>88</sup> we can't claim to have seen biology's Newton yet.

Nevertheless, the case for mechanism is much stronger today than in Kant's time. If, contra Kant, mechanistic explanation of life can't in principle be ruled out as a human possibility, what remains viable in his perspective on function? Perhaps a bit surprisingly, it seems that the conceivability of mechanistic explanation doesn't much affect teleology's role as constitutive of a separate domain of study (life, biology). At most it shows that the indispensability of the teleological maxim cannot justify a dogmatic *a priori* verdict on the question of how far mechanism can reach. The exposition of Kant's view above points to two aspects of teleological attribution in Kant, which may be called the identificatory and the quasi-explanatory. Teleology in its identificatory role singles out the biological object as a functional unit, an organism. The quasi-explanatory use of teleology serves to provide a "law," or at least some order, unifying the vast number of causal chains that interact in the organism in ways otherwise wholly contingent.<sup>89</sup> Successful mechanistic explanation would lessen the need for quasi-explanatory teleological attributions, leaving their identificatory function intact.

If one day mechanistic explanation of the origin of life and the formation of organisms were nearly complete, it would be possible to take functional locutions as unproblematically metaphorical. Even then, we would presumably still need teleology, understood along projectivistic lines, as an identificatory device for delimiting the structures we wish to investigate. That we are not even close to that point is shown by the unreduced teleology that surfaces time and again even in the most austere naturalistic analyses. Indeed, modern molecular biology, though officially naturalistic, makes free use of a

---

<sup>88</sup> See for instance the discussion in Kirschner, Gerhart and Mitchison 2000.

<sup>89</sup> Kant again and again stresses that teleological attribution is reflective and not determinative, so he is clearly aware that the quasi-explanatory use of teleology is not genuinely explanatory.

*Biological Functions in a Kantian Perspective*

quasi-explanatory teleological notion of genetic information.<sup>90</sup> Furthermore, in singling out the genes as the sole locus of informational content, disregarding contributions from other developmental resources, it treats the genome as a *Naturzweck*, thereby arbitrarily displacing the teleological perspective from the organism as a whole in its connection to the environment.<sup>91</sup>

---

<sup>90</sup> See the discussions in Maynard Smith 2000 and Griffiths 2001.

<sup>91</sup> For a critique of this move from a Kantian point of view, see Moss 1992.

# The Antinomy of Teleological Judgment

The second part of the *Critique of the Power of Judgment* (CJ), the *Critique of the Teleological Power of Judgment*, follows the standard Kantian architectonic in that it consists of an *Analytic*, a *Dialectic*, and a *Methodology*. I will here give a textual interpretation of the *Dialectic* (CJ §§ 69–78), focusing on its antinomy. The aim is to understand which principles are supposed to be involved in the antinomy, and how Kant wants to resolve it. In the course of clarifying the overall structure of the *Dialectic*, the role of the complex discussion in §§ 76–77 on the nature of our discursive understanding will also be investigated. Some points will be made in passing concerning the objects that are judged teleologically, namely organisms, as well as on the methodology of biology. Questions concerning the role of the *Dialectic* for the mediation sought in the third *Critique* between theoretical and practical reason, important as they are for Kant's wider project, will not be considered.

There is a certain tension between the aims of explaining, on the one hand, what the antinomy consists in and, on the other hand, how to make sense of the overall composition of the text. A rather natural interpretation of the antinomy and its solution seems to make the main part of the *Dialectic* unnecessarily long and digressive, while on the other hand efforts to get a grip on the *Dialectic* as a whole tend to lead commentators to dismiss Kant's own prescriptions on how to treat the antinomy. The present interpretation will take Kant's stated views on the antinomy and its solution seriously, while explaining his inclusion of seemingly unnecessary sections by showing their roles in his argumentation.



## 1. PRESENTATION OF THE ANTINOMY

In the opening paragraphs of the *Dialectic*, Kant presents an antinomy of the power of judgment. Its place in Kant's architectonic is explained in § 69. Like other Kantian antinomies, it is said to be a conflict that carries with it a necessary illusion dissolvable only by a critique of our cognitive capacities. This conflict gives rise to a natural dialectic, just as the antinomy of reason of the *Critique of Pure Reason* (CPR) does. But it is an antinomy of the power of judgment, not of reason, because the conflicting propositions are maxims for reflective judgment, not principles for determining objects.

Reflective judgment does not determine objects, that is, it does not apply a given concept or law to particulars, but seeks a concept for a given particular. As the concept is not given, reflective judgment has to provide its own principle for its application. This cannot be objective (if we had an objective principle, we would not need reflection here), it can only be a subjective maxim. In the context of judging organized beings, a maxim of final causes is given us by reason, prompted by "particular experiences that bring reason into play" (CJ 386).

What is the other maxim that enters the antinomy? In § 70 it is specified as a maxim of mechanism. That this mechanistic maxim is closely connected to the principle of causality of CPR seems clear, but the maxim is a regulative principle for reflective judgment, and not a constitutive principle determining objects as causality is. The maxim, however, is given *a priori* by the understanding to judgment (CJ 386).<sup>1</sup> Very schematically, three main lines of interpretation of the relationship between causality and mechanism seem available: (i) It is one and the same principle, but Kant has changed his mind on its status, from constitutive in CPR to regulative in CJ.<sup>2</sup> (ii) The principle is the same, but it has a different role in CJ, where it is used reflectively, from what it has in CPR, where it is used for determining

---

<sup>1</sup> What this means will be touched upon in section 8.

<sup>2</sup> See for instance Ernst 1909, 64. Butts (1984, ch. 9) may be seen as a variation on this theme, though he construes Kant as having treated causal mechanism regulatively even in CPR.

objects.<sup>3</sup> (iii) The concept of mechanism differs from that of causality; it adds something to the constitutive core of efficient causality. This further content is not constitutive of objects, but only regulative or useful for reflection. Interpretations of this type can be further subdivided. For Marc-Wogau the mechanistic maxim contains the constitutive principle of causality supplemented by the non-constitutive claim that nothing more is needed for the explanation of nature, and Melnick in a somewhat similar vein construes the mechanistic maxim as a reflective way of eliminating the randomness of contingently occurring events (each of which is causally determined).<sup>4</sup> Admittedly, these interpretations, when spelled out, are rather close to those in (ii), as they consider a special regulative use of the constitutive causal principle to be operative in CJ. A more distinctive representative of (iii) is McLaughlin, who forcefully defends the thesis that whereas constitutive causality involves a determination of time-order, the maxim of mechanism is a further, regulative assumption about the relation of part to whole, according to which a whole is always causally determined by its parts.<sup>5</sup> Still another position in this group is that of Ginsborg,<sup>6</sup> for whom mechanistic explanation involves reduction to “the mere forces of matter as such” or to forces of particular kinds of matter, a stronger requirement than that of merely providing a causal account. I will return to the question of mechanism’s status in section 8.

When we use these two maxims for reflection, they may seem incompatible, so that “a dialectic will result that will make the power of judgment go astray” (CJ 387). The mechanistic maxim is the thesis:

T            *All production of material things and their forms must be judged as possible in accordance with merely mechanical laws.*

The maxim of final causes is the antithesis:

---

<sup>3</sup> Makkreel 1991, 53. Compare Zumbach (1984, 142n 23), who distinguishes two concepts of explanation with the same causal principle used constitutively and regulatively, respectively.

<sup>4</sup> Marc-Wogau 1938, 227, Melnick 1973, 121–130; compare Frost 1906, 321.

<sup>5</sup> McLaughlin 1989, 364; 1990, 152–155. Allison (1991, 27) agrees.

<sup>6</sup> Ginsborg 2001, 244.

*The Antinomy of Teleological Judgment*

A            *Some products of material nature cannot be judged as possible according to merely mechanical laws (judging them requires an entirely different law of causality, namely that of final causes). (CJ 387)*

As an antinomy pertaining to the question of the “production of material things,” it is connected to the biological debate of the time concerning the generation and development of organisms. Kant’s term *Erzeugung* (production) is closely related to *Zeugung* (generation).<sup>7</sup>

Kant immediately presents another pair of principles contradicting each other. These are derived from T and A, but make no reference to our judging. They are formulated as constitutive principles for determining objects.

T’            *All production of material things is possible in accordance with merely mechanical laws.*

A’            *Some production of such things is not possible in accordance with merely mechanical laws.*

These principles contradict each other, so one of them has to be false. This would be an antinomy of reason (though not of reflective judgment) (CJ 387). But none of them can be proved; we cannot know *a priori* how things are produced by empirical laws of nature. So these principles are not genuine *a priori* principles of reason. In the *Critique of Pure Reason*, each of the antinomies consisted of a proven thesis and a proven antithesis. This made reason’s conflict with itself very threatening. If proven propositions contradict each other, there can be no certainty in the use of reason. Kant’s strategy in resolving the antinomies of CPR is to show that they presuppose transcendental realism, the alternative position to his own transcendental idealism.<sup>8</sup> When the assumptions of transcendental realism are replaced by those of transcendental idealism, the contradiction vanishes, and thesis and antithesis can both be true, or both be false, as the case may be.

---

<sup>7</sup> In formulating the theses and antitheses, Kant has *Erzeugung* in all cases, except in the antithesis maxim (A) where he has *Produkte*. I follow Pluhar’s rendering of *Erzeugung* as “production” (see Kant 1987, 267).

<sup>8</sup> See Allison 1983, chs. 2–3.

*The Antinomy of Teleological Judgment*

The relation between the antinomy of principles T' and A' on the one hand, and the antinomy of reason of CPR (and that of the *Critique of Practical Reason*) on the other hand, will be explored later. For the moment I just note that what Kant says about the antinomy in the rest of § 70 and in § 71 seems to indicate that the antinomy he is going to resolve is that between T' and A', and that the resolution consists in seeing that we have no constitutive principles here, but only maxims for reflective judgment, T and A. "All appearance [*Anschein*] of an antinomy between the maxims [...] therefore rests on confusing a fundamental principle of the reflecting with that of the determining power of judgment" (CJ 389). As soon as we clearly understand that we have nothing but maxims for judging (*Beurteilung*) here, not principles determining objects *a priori*, the antinomy is dissolved, because the maxims T and A do not have to be taken as contradictory. The thesis just says that all "products" of nature must be judged according to mechanism, which should be applied "so far as I can," and that is not, according to how Kant presents the matter here, incompatible with the other maxim, which says that we occasionally should reflect on some "products" using a different principle, that of final causes (CJ 387-388).

At this point, it seems that this is all there is to the solution. The antinomy of teleological judgment is a conflict between two putative constitutive principles of reason. The solution consists in pointing out that the principles have the more modest status of being regulative maxims for reflective judgment. This is the interpretation that used to be offered in earlier commentaries,<sup>9</sup> and it is still common.<sup>10</sup> But at least since Marc-Wogau (1938, 225), there have been dissenting voices. According to Marc-Wogau, the genuine antinomy of judgment is that between T and A. In § 69 Kant announces an antinomy of judgment, but after presenting the constitutive principles T' and A' in § 70, he immediately tells us that the contradiction between them would indeed "be an antinomy, though not of the power of judgment, but rather a conflict in the legislation of reason" (CJ 387). So even if the conflict of principles T' and A' is neutralized by their demotion to

---

<sup>9</sup> See for instance Ernst 1909, 64. A number of representatives of this interpretation are referred to in McFarland 1970, 120–121.

<sup>10</sup> See for instance Butts 1984, Düsing 1986, 264–265, Rang 1993 and Effertz 1994.

the status of regulative maxims, there still must be an antinomy between these maxims, if there is to be an antinomy of judgment. This is one reason why many of the more recent commentaries locate the antinomy between the maxims T and A.<sup>11</sup>

It might be useful to list some considerations that favour the view that the real antinomy is between T and A.

1) *Antinomy of judgment.* Kant says, as mentioned above, that the antinomy of constitutive principles isn't an antinomy of judgment; but in § 69 he has promised us an antinomy of judgment. Pertinent to this point is that earlier in the *Critique of the Power of Judgment* (§ 57, Remark II) he says that "there are *three kinds of antinomy* of pure reason" which he lists as concerning the theoretical use of the understanding, the aesthetic use of judgment, and the practical use of reason, respectively (CJ 344–345). The antinomy of teleological judgment is not mentioned in this list, an omission that might be taken to show that it is no "antinomy of pure reason" at all, but rather an antinomy of judgment,<sup>12</sup> whereas the antinomy of reason between T' and A' referred to in § 70 does not count as a genuine antinomy (the reason for which is found in point 6) below).

2) *Too long.* If the antinomy is solved by no longer mistaking regulative principles for constitutive ones, then there is no more to consider after § 71. So why does the *Dialectic* include §§ 72–78?<sup>13</sup>

3) *Preparation.* The title of § 71 is *Preparation for the resolution of the above antinomy*. So it isn't the solution itself.<sup>14</sup>

4) *Contradiction of the maxims.* If the antinomy is between the constitutive principles, and the solution amounts to seeing that they are merely regulative, then, as regulative maxims, they must be compatible. Kant says they are (CJ 387–88), but in fact they are contradictory.<sup>15</sup> To remove *this* contradiction, which amounts to an antinomy between T and A, must be the real task of the *Dialectic*.

---

<sup>11</sup> McFarland 1970, Siewert 1974, Zumbach 1984, 130–131, McLaughlin 1989 and 1990, Allison 1991, Zanetti 1993.

<sup>12</sup> McLaughlin 1990, 124–125.

<sup>13</sup> Marc-Wogau 1938, 215.

<sup>14</sup> Frost 1906, 336, McFarland 1970, 121.

<sup>15</sup> McLaughlin 1990, 134. According to McLaughlin (1990, 139), Hegel was the first commentator to clearly state that the regulative maxims are no less contradictory than their constitutive counterparts.

### *The Antinomy of Teleological Judgment*

5) *Non-necessity*. An antinomy in Kant's sense is not just any arbitrary confusion, but a deep dialectical illusion engendered by reason itself, and removable solely by reason's critique of itself. But confusing regulative with constitutive principles is not a mistake necessitated by the nature of our cognitive faculties.<sup>16</sup>

6) *Lack of proof*. Related to 5) is the fact already mentioned that the antinomy of CPR consists of *proven* theses and antitheses. The propositions T' and A' differ very much from that: "[r]eason can prove neither the one nor the other of these fundamental principles, because we can have no determining principle *a priori* of the possibility of things in accordance with merely empirical laws of nature" (CJ 387). So the conflicting propositions, being entirely unproved, and even unprovable, do not have the status required for an antinomy.

What then is the antinomy of teleological judgment? If it is the conflict between T and A, as points 1)–6) suggest, why does Kant immediately try to take away even the semblance of a conflict between these regulative maxims by claiming that they do not contradict each other? On the other hand, to repeat 1) once again, if the real antinomy is the conflict between T' and A', why has he announced an antinomy of judgment instead of one of reason, which is his own designation of this conflict? Kant's various statements on the antinomy of teleological judgment are difficult to unite in a consistent way; they almost seem to amount to an antinomy of their own.<sup>17</sup>

## 2. PROGRAMME FOR AN INTERPRETATION

The present reading of the *Dialectic* will attempt to make sense of these conflicting tendencies in the text. While it is certainly true that the solution to the antinomy cannot *only* consist in the removal of a pair of spurious constitutive principles and their transformation into regulative maxims, nevertheless this move is Kant's central concern.

---

<sup>16</sup> McLaughlin 1989, 360.

<sup>17</sup> This situation leads Frost (1906, 341) to claim that Kant did not even intend to resolve the antinomy.

### *The Antinomy of Teleological Judgment*

Though T' and A' are not valid constitutive principles, all the same they have a strong semblance of being that, owing to their relations to the principle of causality on the one hand, and to the special properties of organisms on the other. This semblance brings forth a transcendental illusion that to some extent remains also when the principles are understood as regulative maxims. And even assuming the regulativity of the principles, there is still the question of how they can co-exist as maxims. Though Kant claims that T and A do not contradict each other, as they stand they certainly do. Kant has more to do in explaining their consistency than what he offers in § 70. So one task for the discussion in §§ 72–78 is to actually show that we only have regulative maxims for reflective judgment here, not principles for the determination of objects. That Kant states this in §§ 70–71 can be seen as a declaration of his goal, of what he intends to show in the *Dialectic* as a whole. The need to substantiate this claim explains the long discussion of the subsequent paragraphs.

Some comments on points 1)-6) are in order to further delineate the present interpretation.

1) *Antinomy of judgment.* This point shows that there must be more to the antinomy than merely the conflict of T' and A'. About their regulative counterparts, Kant states in § 69 that “between these necessary maxims of the reflecting power of judgment there can be a conflict, hence an antinomy” (CJ 386). Thus we should find an antinomy of some sort also between T and A. This antinomy, I propose, consists in our tendency to mistake the maxims for constitutive principles, *even after* they have been formulated as regulative maxims. It is the difficulty of avoiding a determinative interpretation of the maxims of reflection that lies at the root of the antinomy. As Allison points out, the maxims “seem to involve some kind of ontological commitment,” and this appearance is only removed “by a transcendental critique.”<sup>18</sup>

2)–3) *Too long & Preparation.* It has to be established that our maxims are merely regulative. This has been stated, but it is argued for in §§ 72-78. That is why § 71 is only a preparation to the solution.

4) *Contradiction of the maxims.* The maxims are as contradictory as the constitutive principles, according to this point. This is correct, since the thesis says that all production of material things must be

---

<sup>18</sup> Allison 1991, 31–32.

*The Antinomy of Teleological Judgment*

judged to be possible in terms of merely mechanical laws, whereas the antithesis says that some products of nature cannot be judged to be possible in terms of merely mechanical laws. Yet Kant denies that the maxims contradict each other. Later in his discussion (in § 72) he says that they are “disparate” and can well be united, unlike the constitutive versions of the principles that are contradictorily opposed (CJ 391). However, Kant’s denial that the maxims contradict each other is not, (as has been proposed<sup>19</sup>) based on the claim that maxims, being subjective directives for action, cannot stand in logical relations. That Kant does not consider maxims to stand outside of the domain of logic is clear from his moral philosophy, centred as it is on the logical consistency of universalized maxims. In the present case he tries rather to get round his preliminary formulations, by offering reinterpretations of the maxims. What they really mean, according to Kant, is that we ought always to reflect on natural products in terms of mechanical law, as far as we can, but that we also, on occasion, should consider the matter from the point of view of final causes. What this can mean in biological practice will later be exemplified. For the moment I shall only note that this attempt to disarm the antinomy has some likeness to what Kant says in CPR on how to treat seemingly contradictory regulative maxims. In a discussion of the heuristic regulative maxims of homogeneity, specification and continuity, he points out that the first two of these appear to contradict each other. The maxim of homogeneity (somewhat akin to what Wittgenstein called “a craving for generality”) directs us to search for unification of empirical laws under higher, more general laws. Specification, on the other hand, is an equally valid approach, striving to divide general laws or concepts into more particular ones, and to recognize maximal variation (A642/B670–A668/B696). Kant’s way of treating the opposite tendencies represented by these maxims is relevant for the antinomy of judgment. “If merely regulative principles are considered as constitutive, then as objective principles they can be in conflict; but if one considers them merely as *maxims*, then it is not a true conflict, but it is merely a different interest of reason that causes a divorce between ways of thinking. Reason has in fact only a single unified interest, and the conflict between its maxims is only a variation and a reciprocal limitation of the methods

---

<sup>19</sup> By Butts 1984, 272–273.



satisfying this interest” (A666/B694). At least to some extent this resembles the conflict between T and A. These are also conflicting maxims, treated by Kant as fruitfully complementing each other, rather than as contradictory.<sup>20</sup>

However, a difficulty remains. If this is the justification of Kant’s claim that T and A do not exclude each other, why does an antinomy and an entire *Dialectic* arise for teleological judgment? In CPR there certainly is no antinomy of homogeneity and specification.<sup>21</sup> This discrepancy can be explained on the assumption that the antinomy of judgment largely consists in the *tendency* to interpret maxims T and A as if they implied T’ and A’. This dialectical temptation is not overcome just by stating that the maxims are regulative. Their regulativity has to be argued for. Moreover, it has to be shown that concurrent use of the maxims in judging organisms does not bring with it unwanted ontological commitments leading to contradiction. The urgency for a dissolution of the conflict is greater in this case than in that of the maxims of homogeneity and specification, which concern the hierarchical relations of laws or concepts and can rather easily be conceived of as complementing each other; they were unified by Kant in the maxim of continuity, that brings together the other two by prescribing both an ascent to higher generality and a descent to subordinated concepts, thereby creating a systematic order (CPR A658/B686). Upon observing properties in organisms that seem to call for some sort of holistic or finalistic explanation, while at the same time pursuing the mechanistic programme of science, it is rather more difficult to abstain from taking regulative maxims as constitutive principles for determining organisms ontologically.

---

<sup>20</sup> See also Rang (1993, 69) on complementarity.

<sup>21</sup> An interesting suggestion by Peter (1992, 231–234) is that this is the point where CJ corrects the *Dialectic* of CPR. He takes the occurrence of an antinomy of maxims to show that reason is subordinated to reflective judgment, inasmuch as reason is incapable of resolving the antinomy without the help of an autonomous principle of judgment. The claim in CPR that reason on its own is able to consider conflicting maxims as complementary would accordingly be a mistake, corrected in CJ. It is not possible here to pursue the difficult question on the relation between CPR’s *Dialectic* and the *Critique of the Teleological Power of Judgment*.

*The Antinomy of Teleological Judgment*

5)–6) *Non-necessity & Lack of Proof*. These points stress the artificiality of the antinomy between T' and A'. As unproven, the principles are no threat to reason. Mistaking regulative maxims for constitutive principles is to be sure a grave error, but it does not amount to an antinomy in the sense of the first *Critique*.

To this it can be answered that even if the conflict of these principles fails to satisfy the conditions for an antinomy as presented in CPR, this may not be the only model for antinomies in Kant's philosophy. In the *Critique of Practical Reason*, there is an antinomy between the maxims of virtue and happiness. The concept of the highest good, a necessary "object" or goal of pure practical reason, contains virtue and happiness combined. Since these are not identical concepts, their connection must be synthetic, and Kant takes this to imply that the relation between them is causal. If the highest good is to be possible, then either happiness is the cause of virtue, or virtue is the cause of happiness. The first alternative is absolutely impossible, Kant says, since he has already proven that the desire for happiness has nothing to do with virtue. The second is also impossible, it appears, because happiness depends not on virtue but on events determined by the laws of nature, indifferently to the virtue of those involved.<sup>22</sup> The antinomy is solved by the distinction between appearance (*Erscheinung*) and intelligible world. Considering oneself as belonging to the intelligible world, it is conceivable that one's virtue will cause a proportionate happiness. So the second alternative is not impossible, given transcendental idealism, which allows an empty space in which the intelligible world is at least thinkable. Thus the highest good is practically possible.<sup>23</sup>

Kant's procedure reflects a development in the notion of antinomy.<sup>24</sup> The thesis, that happiness causes virtue, is absolutely false. Thus, we are very far from the proven theses and antitheses of CPR.

---

<sup>22</sup> *Critique of Practical Reason*, Ak. 5:113.

<sup>23</sup> *Critique of Practical Reason*, Ak. 5:114–115. Admittedly this is but one of a number of ways to construe this rather unclear text. See Milz 2002 for a thorough discussion of the alternatives.

<sup>24</sup> Hinske 1965 shows that Kant's use of the term "antinomy" isn't static, and that already in CPR it has three applications: to conflicting propositions, to conflicting laws, and to the state of reason when involved in a dialectic. See also Milz 2002, 204.

### *The Antinomy of Teleological Judgment*

And the solution of the antinomy is not as in CPR, that both propositions are false (mathematical antinomies), or that both may be true (dynamical antinomies). The solution is that the thesis is false and the antithesis possibly true, so that the possibility of the highest good (which requires that the falsity of both statements is not certain) is rescued after all. What makes such an antinomy the business of critical philosophy appears to be (i) that its thesis and antithesis are the only possible ways to construe a concept necessary for Kant's moral philosophy (that of the highest good), and (ii) that a version of the transcendental idealist distinction of appearance and thing in itself is required to avoid the conclusion that this concept is impossible.

The antinomy of the second *Critique*, though not perfectly analogous to the antinomy of judgment, may in some respects be a better model than that of the first *Critique* for understanding it. The antinomy of the constitutively interpreted principles T' and A' is also an opposition between unproven propositions. Their semblance of validity is derived from their possible relevance for explaining the status of the concept of a natural purpose (*Naturzweck*, the problematic Kantian concept of the organism). This is somewhat analogous to (i) in the case of the antinomy of the second *Critique*. Though, according to Kant, the existence of natural purposes is not necessary in the way that the highest good is for moral philosophy, it is nevertheless a necessity for the life sciences (and it also has a larger role to play in the project of providing a transition between theoretical and practical philosophy that is the systematic aim of the *Critique of the Power of Judgment*). And, in analogy to (ii), the solution to the antinomy of judgment requires the transcendental idealist distinction of appearance and thing in itself; for the solution does not only consist in replacing T' and A' with their regulative counterparts T and A, but also in showing how these latter can be regulative, and why it is not necessary to take them as contradictory. For this latter task Kant uses the distinction between nature as appearance and its supersensible basis (as will be discussed below).

In the following sections, the interpretation here outlined will be confronted with the text of §§ 72–78.

### 3. THE DOGMATIC SYSTEMS: §§ 72–73

In these sections Kant carries out a critical overview of how earlier systems of philosophy have treated the question of purposiveness in nature. In the present context it is not so much the details of Kant's exposition that concern us, but rather its place in his general argument.

The discussion begins with the remark that we use the concept of final causes as a guide for observing the features of organized beings, and that nobody denies that the concept has this role. In various systems of philosophy attempts have been made to spell out the metaphysical consequences of the appeal to finality. The systems, presented in § 72, have all proceeded dogmatically, which means that they have disputed over objective principles of the possibility of things, principles "*contradictorily opposed*" (CJ 391). Some think that what may look for us like a different kind of causality (finality), is in fact nothing but mechanistic causality of nature. Others have argued that there is, objectively, purposiveness in nature, so that "the productive capacity of nature in accordance with final causes must be held to be a special kind of causality" (CJ 391). These positions correspond to the thesis and antithesis of the constitutive version of the antinomy. Kant calls them "idealism" and "realism", respectively. These labels, which one can easily read too much into, are in this context indicative of the two possible dogmatic positions concerning purposiveness in organic nature. Idealism reduces what looks like purposiveness to mechanism. According to it, purposiveness is only ideal and so has no objective validity. Realism, on the other hand, accepts the objective validity of purposiveness as a special kind of causality. This is a standard use in Kant of "ideal" and "real" as contrasting terms,<sup>25</sup> and has no pejorative sense,<sup>26</sup> even though Kant criticizes the philosophical positions concerning purposiveness designated with these terms.

---

<sup>25</sup> See for instance CPR A27/B43–A28/B44: "Our expositions accordingly teach the *reality* (i.e., objective validity) of space in regard to everything that can come before us externally as an object, but at the same time the *ideality* of space in regard to things when they are considered in themselves through reason, i.e., without taking account of the constitution of our sensibility."

<sup>26</sup> As claimed in Allison 1991, 33.

### *The Antinomy of Teleological Judgment*

It is important to notice that Kant in § 72 not only explores the dogmatic systems that fall under T' or A', but also shows how the regulative maxim A (saying that judging some products of nature requires a non-mechanistic law of final causes) by a natural dialectic is interpreted as if it implied the constitutive principle A'. Actually, the dialectic starts from “the undisputed maxim”<sup>27</sup> that the concept of final causes is a useful guide to the observation of organized beings. This maxim is all right as long as one does not use it for explaining the ultimate origin of organisms, and together with its mechanistic counterpart it is sufficient for the study of nature. But reason has a “presentiment” that by means of the concept of final causes we could “step beyond nature and even connect it to the highest point in the series of causes” (CJ 390). Thus the dialectic gets started, and metaphysical explanations of purposiveness opposing each other present themselves. This dialectic is natural, inasmuch as it results from reason's tendency to search for something unconditioned above conditioned nature, and because of that it requires *critique* in order to be dissolved, which means that it must be shown that the maxims have their ground in our cognitive faculty.

According to Kant's systematization, idealism and realism concerning purposiveness can each take two forms (hyperphysical and physical respectively), depending on whether God or only matter is admitted in the system. We get the following possibilities:

#### *I Idealism concerning purposiveness*

(i) *Casuality (Kasualität)*, “lifeless matter,” represented by Epicurus and Democritus. Here the purposiveness of organisms is explained by blind chance. Kant sees no merit whatsoever in this view: “thus nothing is explained, not even the illusion in our teleological judgments” (CJ 393).

(ii) *Fatality*, “lifeless God,” associated with Spinoza. A supersensible principle is used to explain purposiveness, but this principle is not conceived as intentional. All things are regarded as accidents that inhere in an original being by the necessity of its nature. This view gives a unity to things, but it does not provide any explanation for

---

<sup>27</sup> CJ 390, Pluhar's translation.

purposive features, and it cannot even differentiate purposive things from other things (CJ 393-94).<sup>28</sup>

## II *Realism concerning purposiveness*

(i) *Hylozoism*, “living matter.” The notion of living matter is contradictory, since inertia, lifelessness, is the essential character of matter (CJ 395). According to Kant, life is the power to act on desire.<sup>29</sup> If this power were accorded to matter, natural science would be impossible.<sup>30</sup> The only way to conceive hylozoism is as matter endowed with life (*belebte Materie*), rather than as living matter (*lebende Materie*), so that nature can be thought of in analogy to a living animal.<sup>31</sup> But if we try to explain organized beings by reference to the life of matter, this life in its turn can only be explained by reference to organized beings, from which our acquaintance with such life stems. Our explanation moves in a circle (CJ 394–95).<sup>32</sup>

(ii) *Theism*, “living God.” In this system realism of purposiveness is explained by the intentional causality of an original being. Though this would avoid the shortcomings of the other systems, it cannot be determinately proven. Only by strictly proving that organisms cannot

---

<sup>28</sup> I cannot here present more than a sketch of the discussion of Spinozism. For divergent assessments of Kant’s objections to Spinoza, see Allison 1980 and Zammito 1992, ch. 12.

<sup>29</sup> *Critique of Practical Reason*, Ak. 5:9.

<sup>30</sup> See also Ak. 4:544 (Kant 2002, 252): “The possibility of a proper natural science rests entirely and completely on the law of inertia (along with that of the persistence of substance). The opposite of this, and thus also the death of all natural philosophy, would be *hylozoism*.” For a broader discussion of Kant’s use of the concept of life, see Makkreel 1990, ch. 5.

<sup>31</sup> On Kant’s later discussions on the notion of a world-soul in the *Opus postumum*, see Düsing 1986.

<sup>32</sup> Peter (1992, 227) relates hylozoism to the constitutively interpreted thesis of mechanism (what I have called T’). This has some plausibility if life is considered as a property of matter which, like other material properties, is subjected to causal mechanism. But hylozoism as discussed by Kant is rather the view that matter is endowed with an intentional power to act. As a realism of purposiveness, it is committed to explaining purposiveness in nature as intentional, that is, as due to a special, non-mechanical kind of causality (see CJ 391). That idealism is connected to T’ and realism to A’ seems to be recognized by Peter elsewhere (1992, 223).

be produced by the mere mechanism of nature would we be entitled to postulate determinately that the ground of nature's purposiveness lies beyond nature. But the character of our cognitive power is such that we have no insight into the inner ground of this mechanism. Theism can therefore at most be a subjectively justified point of view, but it cannot be objectively asserted (CJ 395).

Kant's conclusion of his survey of the dogmatic systems is that none of them accomplishes the task of explaining purposiveness in nature. The arguments against the systems presuppose that the purposiveness to be explained has the character of a *Naturzweck*. This concept is that of a purposive unit whose parts are determined by the whole. It might be doubted that the idealists of purpose are committed to this complex concept when they accept (as it is presupposed that all the systems do) that we need final causes as a guide for judging organisms. Kant, on the other hand, apparently considers his account of *Naturzwecke* to be a phenomenological description of what is involved in making judgments about organisms.

From a modern point of view, it could be argued that the system of Casuality loses some of its randomness if supplemented by the theory of natural selection. Perhaps Kant would claim that as natural selection acts on organisms which are conceived of teleologically, as functional entities, purposiveness is taken for granted rather than explained even on this theory.<sup>33</sup>

#### 4. THE CONCEPT OF NATURAL PURPOSE IS NOT OBJECTIVE: §§ 74–75

To show in a more general way that the maxims of mechanism and purposiveness are not principles for the objective determination of organisms, Kant examines in the next sections (§§ 74–75) why it is not possible to use the concept of a natural purpose (*Naturzweck*) dogmatically, that is, for determinative judgment. Such use of the

---

<sup>33</sup> The persistent difficulties for attempts to naturalize the notion of biological function may be an indication that Kant's discussion is of more than merely historical interest. See the preceding essay.

concept would require that its *objective reality* was secured, which amounts to establishing that an object conforming to it is possible (CJ 396, CPR A155/B194). This could be done either by showing that it is an empirical concept, abstracted from experience and therefore unproblematically referring to its object, or by demonstrating that the concept is necessary for the possibility of experience (as in the case of the categories). The second alternative is out of the question since the concept is “empirically conditioned,” dependent upon “certain conditions given in experience” (CJ 396), namely the existence of organisms, and it might therefore seem to be an ordinary empirical concept. But it is not, since it contains the concept of a causality through purposes, and we do not know any purpose or intention that produces organisms (compare CJ 374). The problem is not to establish the objective reality of the concept of a causality through purpose: any work of art does that. It is instead the amalgamation of a causality through purposes independent of human intention with causality of nature in one and the same thing that is problematic. The concept of a natural end “includes natural necessity,” but on the other hand it refers to something “supersensible,” so the concept is transcendent or “excessive” for determinative judgment though immanent for reflective judgment (CJ 396). Purposiveness of nature is therefore a “subjective fundamental principle merely for the reflecting power of judgment, hence a maxim that reason prescribes to it” (CJ 398). In § 75 the consequences of this result are discussed. If we had to say objectively (judge determinatively) that some things of nature are produced intentionally, we would affirm the existence of the God of the theistic system. But used as a regulative maxim, on the other hand, this purposiveness of nature carries no ontological implication, holding only from a “human” point of view (CJ 400).

As for the mechanistic principle, it is insufficient for explaining the structure and the generation of organized beings (CJ 400). Why is this so? The answer depends on what Kant more precisely means by “mechanical,” and a discussion of that will have to wait until the presentation of the general outline of the *Dialectic* is finished. Let me just note for the moment that one line of argument for the insufficiency of mechanism is that the organism is contingent relative to the transcendental laws constitutive of nature as a physical system. This contingency, which amounts to a surplus of order unexplainable in mechanistic terms, appears to require an explanation. On the other



hand, we cannot assert that in nature “there *could* lie hidden no ground sufficient for the possibility of organized beings without the assumption of an intention underlying their generation” (CJ 400). We cannot know if organized beings ultimately are mechanically produced, but nonetheless we know that the principle of mechanism is inadequate to account for their production. The possibility to combine the two claims – that the mechanistic principle cannot provide the required explanation, and that organisms may originate mechanically through some hidden capacity of nature – is presumably connected to the regulativity of the mechanistic principle. It is a paradoxical claim that beings which perhaps are mechanically produced can at no event be explained mechanically, but the paradox is somewhat softened by noting that the thought of the production of organisms concerns an ontological (and for us inaccessible) mechanism, whereas the attempted explanation is mechanistic in an epistemological sense. In § 71 Kant asserts explicitly that the inability of mere mechanism to provide an explanatory ground for the production of organized beings holds “with respect to our cognitive faculty” (CJ 389), which amounts to saying that it is a matter of reflective judgment (compare CJ 395: “we deal with [a concept] merely critically if we consider it only in relation to our cognitive faculties”, and such critical treatment “is lawful merely for the reflecting power of judgment”).

## 5. THE PECULIARITY OF THE HUMAN UNDERSTANDING: §§ 76–77

At this point in his exposition, Kant embarks on a digression touching upon some of the fundamentals of transcendental philosophy, intended “for elucidation” of the issue that concerns us, rather than as a proof (CJ 401). “Perhaps never so many deep thoughts have been compressed in so few pages as in [CJ] § 76,” Schelling wrote in 1795.<sup>34</sup> This *Betrachtung* (§ 76) is followed by a section that connects these most general Kantian tenets to the question of judgments about

---

<sup>34</sup> Schelling 1907 [1795], 94.

natural purposes (§ 77). The present discussion will focus on what I consider important for the overall argument of the *Dialectic*, without any presumption of covering all the questions that these sections raise.

In a survey of our mode of cognizing, Kant picks up the distinction between the modal concepts possibility and actuality (*Wirklichkeit*) (CJ 401–402). He locates the ground for this distinction in the nature of our cognitive faculties. As we need two separate components for cognition, namely the understanding providing concepts and sensibility providing intuition, there is room for distinguishing possibility from actuality. Concepts concern the possible because they are general predicates that possibly apply to an object, though on their own they can never present an actual object for us, whereas intuitions actually give us objects (though these cannot be cognized without the application of concepts). Possibility is the “position [*Position*] of the representation of a thing with respect to our concept and, in general, our faculty for thinking,” while actuality is the “positing [*Setzung*] of the thing in itself (apart from this concept)”<sup>35</sup> (CJ 402). The very possibility of a distinction between possibility and actuality is thus a consequence of the twofold nature of our cognitive faculty. This is in line with some basic tenets of the *Critique of Pure Reason*.<sup>36</sup> As in CPR (see for instance B145), here too Kant deploys the notion of a different understanding contrasting to the one we have. Our understanding is discursive, in need of data given to it from sensibility (as a receptive faculty), on which to apply its concepts. A non-discursive, intuitive understanding “would have no objects except what is actual,” since it would intuit without sensible intuition and need no concepts (CJ 402). Its thinking would thus be an immediate intuiting; hence the notion of an intuitive understanding is the same as the notion of an intellectual intuition used in CPR (for instance at B72). While the intuitive understanding is conceived as a radically different kind of understanding, an

---

<sup>35</sup> The phrase “thing in itself” should not be taken here in the sense of *noumenon*, but rather in what Meerbote 1974 calls the empirical respect, where it means an ordinary thing.

<sup>36</sup> See A50/B74–A51/B75 for a characterization of the two sources of cognition, and A218/B265–A226/B274 for an account of possibility and actuality.

intellectual intuition would be an alternative faculty of intuiting, spontaneous instead of receptive and therefore non-sensible, an *intuitus originarius* in contradistinction to our *intuitus derivativus* (B72). That the notions of intellectual intuition and intuitive understanding are nonetheless identical can be seen in that both are ideas of a unitary faculty which immediately produces its object, a faculty for which the distinction between concept and intuition, as well as that between possibility and actuality, is cancelled. That the two designations refer to the same idea can also be seen by reflecting upon what to call the cognition possessed by an intuitive understanding. An *intuitive understanding*, a faculty of understanding that intuits, has intuitions that are intellectual and not sensible, since it needs no recourse to a separate receptive faculty, that is, it has *intellectual intuition*.

Kant uses this notion as a contrast to our discursive understanding, in order to direct our attention to the peculiarity (*Eigentümlichkeit*) (CPR B145, CJ 405) of the latter. In CPR the function of the idea of an intellectual intuition is to indicate that for our intuition objects have the status of appearances, structured *a priori* by features belonging to our special mode of intuiting (space and time). Here, a similar move is made to show that the purposiveness we attribute to some things of nature has its ground in a peculiar feature essential to our cognitive faculty, and is therefore not objectively ascribable to these things. There is a difficulty, however, concerning the application of this strategy here (even as used for elucidation rather than proof), because the claims of CPR and CJ that the contrast to a unitary cognitive faculty is supposed to elucidate are fundamentally different. In CPR the idea of an alternative kind of cognitive faculty is employed to show that there are constitutive features of our sensibility and understanding that determine appearance *a priori*, whilst in CJ the idea is used in order to establish that the principle of purposiveness necessary for us in the study of organized beings has its root in our peculiar cognitive faculty and not in things considered in themselves, but this time with the consequence that the principle is regulative and not constitutive for experience. In one case the strategy establishes constitutive conditions for experience, but in another case it only establishes regulativity. I will return to this disanalogy in the last section. For the moment we must take a closer look at how an

examination of the discursivity of our understanding is supposed to explain why we have to view organisms purposively.

For an intuitive understanding, the lack of a distinction between possibility and actuality would mean that whatever object it cognizes just *is*, without any difference between what exists and what might exist (CJ 403). The realm of possibility would be entirely absorbed by what is actual, and for that reason the notion of contingency would also lose its significance. Everything that is possible to cognize for such an understanding is already actual, so there is no place for contingency in the sense of possibility of not existing. With the loss of contingency, the corresponding concept of necessity disappears too, leaving the intuitive understanding entirely without modal categories. In Longuenesse's words, this amounts to a "collapse of the modal categories."<sup>37</sup>

In the practical perspective, collapsing the distinction between necessity and contingency brings forth the idea of a perfect being, entirely rational and without sensibility, for which there is no possible lack of fit between the objective necessity for an action and what it actually does. For a being of this kind there would be no moral obligation, since obligation presupposes the possibility not to comply with duty, a possibility excluded in this case.<sup>38</sup> Such a being would inhabit "an intelligible world that harmonized throughout with the moral law" (CJ 403).<sup>39</sup>

In the theoretical as well as in the practical perspective, Kant brings these considerations into connection with the question of the nature of a regulative principle. In fact, his remarks on the intuitive understanding and the intelligible world are presented as examples of cases where the understanding "cannot keep up with" reason (CJ 401). The ideas of reason transcend experience, but the understanding uses them regulatively for the sake of experience by restricting them to the condition that they are valid only with respect to our cognitive faculty, without thereby asserting their validity with respect to the object. In the theoretical case, reason demands the assumption of an absolutely necessary being as "original ground" (*Urgrund*), an idea that is transcendent for the understanding and therefore problematical,

---

<sup>37</sup> Longuenesse 2000, 280n 26.

<sup>38</sup> Compare *Groundwork*, Ak. 4:439.

<sup>39</sup> Pluhar's translation.

### *The Antinomy of Teleological Judgment*

but which, when used regulatively and not for determining objects, is “immanent and secure [...] and appropriate for the human point of view” (CJ 403). In the practical case, the idea of an intelligible world with freedom as its formal condition transcends experience, but it serves as a regulative principle which makes rules for action conforming to this idea valid as moral commands for us (CJ 404).

These examples of ideas of reason transcending the grasp of the understanding and thereby providing regulative principles are then compared to the case of purposiveness as a regulative principle for judging nature. The source of this regulative principle is a peculiar feature of our understanding: that it must proceed from the universal to the particular, which makes the particular contingent with respect to the universal (that is, the conceptual determination provided by the understanding) (CJ 404). Our understanding, operating with concepts, can only determine general characteristics of a given object. That the particular object is given in intuition as existing depends on sensibility as receptivity, so the contingency of the particular is a result of the twofold nature of our cognitive faculty. Reason, on the other hand, requiring complete unification of what our understanding is unable to derive from its universal laws, provides the concept of purposiveness to bring under a law that which is contingent for the understanding. While this concept of purposiveness transcends the reach of the understanding, it nonetheless has a use as a regulative principle for judging particular objects.

The contrast presupposed in this concluding part of § 76, and used to indicate the peculiarity of our understanding, is the idea of an understanding that wouldn't have to proceed from universal to particular, but could determine the particular directly in the universal, without resorting to a regulative principle. In § 77 Kant provides further characterization of this envisaged alternative to our discursive understanding, and of its bearing on the concept of a natural purpose. This concept contains the idea of purpose, which cannot be a principle of determinative judgment since it transcends experience, but is used as a principle for reflecting on some objects of experience (CJ 405). It is thus a regulative idea, and, in accordance with the examples in § 76, identifying such an idea amounts to framing a cognitive faculty different from ours for which there would be no need for the principle. Otherwise, if no such alternative could be conceived of without contradiction, the principle in question could

*The Antinomy of Teleological Judgment*

not be said to pertain to our cognitive faculty (and thus to be regulative) rather than to the object, since the principle would be mandatory for any cognitive faculty whatsoever. It would then be independent from any subjectivity, and hence objective. So we must find some peculiarity in our cognitive faculty that is not necessarily present in every cognitive faculty, enabling us to explain how purposiveness (as part of the concept of natural purpose) can be necessary for us in judging some objects and yet not have more than regulative status (CJ 406). The peculiarity, as already indicated, is discursivity, the dependence of our understanding on a receptivity providing it with intuitions on which to apply its concepts. We thus have to presuppose a non-discursive understanding:

But since intuition also belongs to cognition, and a faculty of a *complete spontaneity of intuition* would be a cognitive faculty distinct and completely independent from sensibility, and thus an understanding in the most general sense of the term, one can thus also conceive of an *intuitive* understanding (negatively, namely merely as not discursive) which does not go from the universal to the particular and thus to the individual (through concepts), and for which that contingency of the agreement of nature in its products in accordance with *particular* laws for the understanding, which makes it so difficult for ours to bring the manifold of these to the unity of cognition, is not encountered. (CJ 406)

Kant here repeats a point explained more fully in the *Introduction* to CJ, namely that the laws given to nature *a priori* by the understanding (the *Grundsätze* of CPR) only determine the possibility of nature in general, leaving the diverse forms of nature undetermined, “contingent as far as *our* understanding can see,”<sup>40</sup> though we still need to judge that these laws are united by virtue of a principle, even if it is not a principle belonging to our understanding. This demand for unification of the particular is expressed in the regulative principle of purposiveness, as an *a priori* principle for reflective judgment, according to which we must conceive of nature as if it were ordered for the sake of our cognitive faculty by a higher understanding. Now, as Guyer has remarked, if the argument of § 77 merely reiterates this

---

<sup>40</sup> CJ 180, Pluhar’s translation.

point, it “would appear to apply to laws of geology and mineralogy just as well as to laws of biology,”<sup>41</sup> and it would thus contribute nothing to account for the specific ascription of purposiveness to organisms implied by the concept of natural purpose. Expecting Kant to address problems particular for organisms, we should therefore look for more than this general point.

That Kant has something more to offer is shown by his further description of the intuitive understanding as an ability to reach the particular from a *synthetic universal* (an intuition of a whole as such), in contrast to the human understanding which has to start from an *analytic universal* (a concept) (CJ 407).<sup>42</sup> The discussion is brought to bear on the question of the relation of part and whole, which is more specifically pertinent to Kant’s notion of the organism as a holistic causal structure, a *Naturzweck*.<sup>43</sup> Two ways to proceed in cognizing an object can be conceived of, a holistic and a discursive, respectively. The holistic procedure, characteristic of an intuitive understanding, goes immediately from whole to parts, without any contingency in the relation of these parts to the whole. The discursive procedure peculiar to our understanding goes in the opposite direction, from parts to whole. This is because the discursive understanding can conceive a natural object as a whole (*ein reales Ganze der Natur*) only by the use of a concept, namely causality, by which it conceives the parts as general grounds<sup>44</sup> (*allgemein-gedachten Gründen*) constituting the whole as effect of the motive forces of the parts (CJ 407). The discursive understanding might thus be called *meristic*, in contrast to the holistic intuitive understanding.

New problems emerge from this unexpected twist in Kant’s discussion. How can we justify the move from discussing the relation between concept and particular to discussing the relation between part and whole? And how can Kant describe our understanding’s handling

---

<sup>41</sup> Guyer 2001, 272.

<sup>42</sup> See also *Opus postumum* (Ak. 22:81) on these contrasting forms of universality.

<sup>43</sup> See CJ §§ 64–65.

<sup>44</sup> As McLaughlin points out (1990, 164n) it is the part that takes the place of the concept and the whole that corresponds to the particular when Kant reworks his distinction of contrasting kinds of understanding in terms of part-whole relations.

of wholes as a conceptual procedure solely in terms of causality? To begin with this second question, we are after all quite capable of immediately seeing a house, for instance, as a whole, without any need for considering the causal contributions of its parts to the constitution of the whole. So our cognitive system, one might object, has the desired holistic capacity, and hence it can't be a feature unique for an intuitive understanding. The reply to the objection is that Kant does not necessarily have to deny the possibility of holistic features in perception, but he would have to ascribe them to sensibility rather than to the understanding, whereas the present discussion is about a peculiarity of our understanding (if also with respect to its interplay with sensibility). To the extent that intuition immediately views the house as a whole, this is not yet cognition. To obtain cognition a conceptual synthesis is needed. It is hard to tell what status intuitions without concepts actually have. Kant famously calls them *blind* (CPR A51/B75), but that does not rule out that they can somehow be felt; it only rules out that intuition on its own yields cognition. In the present context, the question concerns the possibility of cognition of a real whole in a holistic way, so we need not pursue all questions as to non-cognitive (or pre-conceptual) consciousness in Kant. Perhaps it could be said that what we are to conceive in framing the idea of an intuitive understanding is an understanding analogous to our intuition in what regards immediacy, one that (in contrast to our understanding) does not synthesize the manifold in a piecemeal fashion, but cognizes it immediately, as our intuition might be said to feel (but not cognize) it. The intuitive understanding with its synthetic universal would not need conceptual synthesis to avoid being blind.<sup>45</sup>

From this explication of what discursivity amounts to, we can approach the question of why causality is singled out as its *modus operandi*. Why is the peculiarity that prevents our understanding from obtaining a holistic grasp of an object identified as the need to construct the object in terms of the causal contributions of the parts,

---

<sup>45</sup> It is somewhat surprising that Kant chooses to call the intuitive understanding's immediate intuition a *synthetic* universal, considering that usually "synthetic" is linked to "synthesis," and synthesis is precisely the business of the discursive understanding. The reason for his choice is perhaps that a concept (as the result of intellectual synthesis) is an analytical unit (in virtue of its being decomposable in analysis).



rather than as any kind of proceeding from part to whole, causal as well as non-causal? The emphasis on causality is presumably to be explained by the nature of the object under investigation. In introducing the notions of analytic and synthetic universality Kant refers to the special case of production as a causal process: "Our understanding [...] has the property that in its cognition, e.g. of the cause of a product, it must go from the *analytic universal* (of concepts) to the particular (of the given empirical intuition)." (CJ 407). The holistic character that interests us in the organism includes not merely its overall physical form, but even more its causal structure with reciprocal connections from part to whole and from whole to part (CJ 373), distinguishing organisms from other physical objects. We search to understand this structure, and in doing so we apply the concept of causality in the piecemeal fashion, going from one part to the next, that is characteristic of our discursive understanding (a procedure the insufficiency of which necessitates the introduction of the concept of purpose, as we will see below).

Let us now look at the first question raised above: how Kant can pass from a discussion of the relation between concept and particular to one about the relation between part and whole. What has just been said about the meristic procedure of our discursive understanding gives some indication of the connection between these relations. The discursive understanding has to apply concepts to intuitions independently brought to it; thus intuition with the diversity it contains is contingent with respect to the concept. The analytic universal (a concept) can determine the particular thing, but when this particular is a whole with causal relations to its parts, the understanding has to determine the parts conceptually to be able to grasp what they are, and the whole is then understood as produced from the influences of these parts, which are now thought through general concepts. The parts cannot be understood directly from a conceptual determination of the whole, because this leaves their particular configuration undetermined, and hence no holistic causal structure can be grasped.

So when Kant shifts from considering the differing relations between concept and particular in the two kinds of understanding to describing a correlative difference in their way of proceeding from

part to whole, he is not simply confusing these relations.<sup>46</sup> What he is claiming is that discursivity, as a peculiarity not necessarily shared by every conceivable understanding, carries with it a peculiar way of handling part-whole relations (at least with respect to causality), namely merism. This claim is more specific than the general point that we can't derive the diversity of empirical laws from the set of general transcendental laws, and when it is considered in conjunction with the analysis of the organism as showing a special causal structure, it goes some way to answering Guyer's objection, according to which Kant's argument in §§ 76–77 is not enough to single out the judging of organisms as in any way different from that of any other object. Though the discussion in these paragraphs certainly is relevant for the general characterization of the principle of reflective judgment used in all judgments of empirical objects, it also takes on a further question, namely how to handle the amazing internal organization in some products of nature. That the organism (natural purpose) is a concern of § 77 is shown by its heading: *On the Peculiarity of the Human Understanding That Makes the Concept of a Natural Purpose Possible for Us*.<sup>47</sup>

Though Kant is not simply confusing the relations of general to particular and whole to part, Rang argues that he commits a *non sequitur* akin to such a confusion.<sup>48</sup> What Rang detects is an equivocation in the use of the term *part* in Kant's exposition. When Kant claims that the discursive human understanding must understand a whole as dependent on its parts, he speaks of conceptual parts, *Teilbegriffe*. Proceeding from the analytic universal, a thing is understood as a combination of properties (it is red, solid, etc.). On the other hand, when Kant says that the organism displays a causal structure in which the parts are dependent on the whole, the parts in question clearly are material parts. So his conclusion, that the discursive understanding is incapable of understanding the parts of

---

<sup>46</sup> As claimed by Driesch 1924, 369. See also Spindler 1925 criticizing Driesch.

<sup>47</sup> Pluhar's translation (which consistently has "peculiar" and "peculiarity" for "*Eigentümlich*" and "*Eigentümlichkeit*").

<sup>48</sup> Rang 1993, 62–67.

the organism immediately from the whole, depends on confusing conceptual parts and material parts.<sup>49</sup>

Rang similarly criticizes Düsing's attempt<sup>50</sup> to reconstruct Kant's argument. According to Düsing, Kant is justified in considering the conceptual parts with which the understanding operates as corresponding to material parts, because the material parts have to be grasped conceptually. The moving forces of the parts are always conceptually understood, and the whole that they bring forth is considered as the effect of these parts.<sup>51</sup> Rang sees in this the same confusion of conceptual and material parts, since nothing in the discursive mode of cognition precludes that parts as spatial and material can be understood as dependent on a whole, even if the whole, as conceptually determined, consists of its conceptual parts.<sup>52</sup> Furthermore, Rang claims that the Kantian account of space is incompatible with the characterization in § 77 of our discursive understanding, according to which it goes from part to whole and grasps the parts as conceptual.<sup>53</sup> Space is not composed of parts preceding the whole, it is rather a "single all-encompassing" whole whose parts are in it (CPR A25/B39). And these parts are, like the all-encompassing space, intuitional not conceptual, therefore not subject to the structure of the analytic universal.<sup>54</sup>

This last point can be countered by noting that Kant doesn't view space as an ordinary empirical object: "Space is merely the form of outer intuition (formal intuition), but not an actual object that can be outwardly intuited" (CPR B457n). The dependence of spatial parts on space as a whole is therefore irrelevant for the question of our cognition of a material whole (such as an organism). In § 77 Kant compares the unity of space with the "real ground we are seeking" in the context of the wholeness of the organized being, and notes that

---

<sup>49</sup> Rang 1993, 64.

<sup>50</sup> Düsing 1986, 86–99.

<sup>51</sup> Düsing 1986, 65.

<sup>52</sup> Rang 1993, 66.

<sup>53</sup> Rang 1993, 65n. Heyse (1927, 115) also adduces Kant's own views on space as a counter-example to the claim that our cognition is limited to the analytic universal.

<sup>54</sup> Rang 1993, 65n.

*The Antinomy of Teleological Judgment*

“space is not a real ground of products but only their formal condition” (CJ 409).

But there is a further consideration that may serve to clarify two questions: on the one hand why Kant moves from the inability of the discursive understanding to grasp a particular without conceptual determination to a corresponding inability to understand a whole except as the effect of the parts, and on the other hand wherein the mechanical inexplicability of the organism consists. Rang, like McLaughlin (1989, 364) and Allison (1991, 27), considers the real whole to be a spatial whole. While this is certainly one aspect of Kant’s *Naturzwecke*, it should also be stressed that they are temporally extended wholes. According to the analysis of the concept of natural purpose in the *Analytic of the Critique of the Teleological Power of Judgment*, firstly the parts are possible only through the whole, and secondly the parts must also reciprocally bring forth each other as to their “form and combination” (CJ 373). The first condition is not enough, since a thing thus produced might be a work of art. The second condition ensures that the parts organize themselves into the whole, without the need of an independent intentional agent in control of the process. Together the conditions combine efficient causes with final causes, or real and ideal causes as Kant proposes to call them (CJ 373).

An object obeying these conditions would be holistic not merely in a spatial sense (the parts, synchronically, having to be explained with reference to the whole), but also in a temporal sense (the parts, diachronically, generating a whole which is such that it can only be judged to be the cause of the combination of these very parts). The fact that the antinomy is presented in terms of the production (*Erzeugung*) of things also points to organization as a temporal process, with the whole as an end that is futural with respect to the activities of the parts. To be sure, such a final causality is incomprehensible for us; it is incompatible with the Kantian as well as with other accounts of causality. But empirical observation of organisms seems to involve this strange conception of a temporal whole. This prompts the notion of a designed purposiveness, serving as a comprehensible substitute for the unintelligible final causes. If organisms phenomenologically (using the term in a modern, unKantian way) did not manifest temporally and spatially holistic features (concerning the ultimate nature of which Kant makes no

ontological claim) there would be no reason not to employ exclusively the ordinary mechanistic procedures of explanation, and there would then be no antinomy of teleological judgment. Though finalism is incoherent on Kant's view, nevertheless the phenomenological observation of organisms points to features about them that suggest the presence of some sort of backward causation.<sup>55</sup> The question is how to avoid drawing unwanted ontological conclusions from these features.

Here the conception of an intuitive understanding is helpful. In contrast to what is the case for us, the whole available for the intuitive understanding is one in which there is "no *contingency* in the combination of the parts" (CJ 407). Since such an understanding would have no need for a separate source of intuition, neither would it require forms of intuition, such as space and time. We cannot have a positive conception of this, but if we proceed on Kant's assumption that the thought of an intuitive understanding is not self-contradictory, we find that the elimination of time thus conceived eliminates the most problematic aspect of the natural purpose: its suggestion of final causality or reversed time-order. The temporal parts of the organism are, for such an understanding, apprehended as immediately united. In this way we can see that purposiveness, as a concept elicited by the organized being's seemingly deviant causal structure, characterized in terms of causality going both forwards and backwards<sup>56</sup> and clearly incompatible with the principle of causality,<sup>57</sup> is not a feature that has to be objectively (constitutively) attributed to nature. It is now possible to consider it as pertaining to a peculiarity of our understanding (its discursivity), and therefore as only subjectively (regulatively) valid. The discursive understanding, constrained to the linear time-order of the principle of causality, can only view organized beings mechanistically, regarding the whole (the complete organism) as contingently constituted by parts forming each other as well as the whole; but it is not impossible that a different

---

<sup>55</sup> See McLaughlin (1990, 49) for a contrary view.

<sup>56</sup> CJ 372: "*sowohl abwärts als aufwärts*", literally: downwards and upwards.

<sup>57</sup> "The causal nexus, insofar as it is conceived merely by the understanding, is a connection that constitutes a series (of causes and effects) that is always descending [*abwärts*]" (CJ 372).

understanding could cognize this in a way that would leave no room for any contingency. Thus mechanism too can be considered a regulative principle, since it is not valid for any understanding whatsoever.<sup>58</sup>

A second aim of § 77 is to demonstrate why we must use precisely the concept of purpose (*Zweck*) when confronted with organized beings. This too Kant explains as a consequence of the same peculiarity of our understanding, its discursivity. Since the contingency of the parts constituting the organized being appears to be far too great in relation to the whole to qualify for a mechanical explanation, we need another way to explain it, one that does not oppose the understanding's causal principle by bringing in some alternative form of causality. This we find in the concept of intentional purposive production. With this concept, final causes are tamed by being attributed to an intelligence, fashioning the end-product out of an antecedent conception of the whole. This is the concept of purpose (*Zweck*): "purpose is the object of a concept insofar as we regard this concept as the object's cause (the real ground of its possibility)" (CJ 220). The peculiarity of the understanding, its having to cognize a whole as the joint effect of conceptually determined parts (§ 77), makes it insufficient for the explanation of organized beings, and therefore the power of judgment is forced to use the idea of purpose. By appealing to a representation of the whole to be produced (a representation present already from the start), the discursive understanding finds a way of combining its mechanistic mode of thinking with the idea of a determination of the parts by the whole. "Hence such a whole would be an effect, a product, the representation of which is regarded as the cause that makes the product possible" (CJ 408). Appealing to purpose is the only way compatible with discursivity for the understanding to conceive of the whole as prior to the combination of the parts.<sup>59</sup>

Thus, both mechanism and teleology (the appeal to final causes on the analogy of artistic production) are to be assigned the status of regulative maxims for judging organized beings, and the need for these principles has been traced to its transcendental source. Thereby an explanation has also been given of the depth of the inclination to

---

<sup>58</sup> Compare the account in Zumbach 1984 (especially p. 135).

<sup>59</sup> Compare Düsing 1986, 96.

treat the principles as constitutive. If they are rooted in human discursivity, they are naturally taken as constitutive; but reflection upon their origin provides a critique, so that even though the illusion of objectivity “still deceives, it does not defraud,” to borrow a phrase from the *Critique of Pure Reason* (A422/B450).

We tend to interpret the two relevant principles for investigating organized beings (the mechanistic and the teleological) as expressing opposing properties ascribed to these beings, but for a different understanding the parts would be directly determined by the whole, so for it there would be no difference (and no conflict) between the principles.<sup>60</sup> The demotion of the maxims to regulative status is part of the solution of the antinomy. What Kant still needs to do is to show how T and A, as maxims for reflective judgment, can be treated also by us (and not only as conceived from the vantage point of a different understanding) as complementary rather than as contradicting each other.

## 6. UNIFICATION OF THE PRINCIPLES: § 78

Kant’s resolution of the antinomy, in accordance with his usual way of handling antinomies, rests on an appeal to a possible solution of the conflict in “the supersensible” (CJ 412). It is tempting to assimilate this manoeuvre to the strategy of the third antinomy in the *Critique of Pure Reason*. That antinomy, in a very rough and compressed formulation, consists in the possibility of proving (by means of two *reductio* arguments) that the assumption of the exclusivity of the causal law leads to the result that there are uncaused causes (freedom), but that such an uncaused cause is impossible because it contradicts the concept of a causal order of nature. The solution consists in showing that the universal causal law and the possibility of uncaused causes are not opposed if we consider them as applying to an object in different respects (*Beziehungen*) (CPR A536/B564), namely as appearance (to which the causal law applies) and as thing in itself (for which freedom is at least not impossible). Transposed to the antinomy of teleological judgment, the

---

<sup>60</sup> Compare Marc-Wogau 1938, 216.

*The Antinomy of Teleological Judgment*

structure of this solution would correspond to the result that mechanism is valid for appearance while teleology has its possible validity in a thing in itself-perspective. Such a reading is offered by Guyer, according to whom “a tension between mechanical and final explanation inevitably suggests itself, which is ultimately resolved by the suggestion that we conceive of these two forms of explanation as applying at two different levels, namely, the sensible and the supersensible.”<sup>61</sup>

This is not Kant’s way of handling this antinomy, however. The crucial step of § 77, bringing in the conception of an intuitive understanding, showed that mechanism and final causes do not necessarily differ for every possible understanding. It also explained why we have to conceive of final causes in the form of a representation of the whole as cause of the whole, what Kant terms “purpose” (*Zweck*). Purpose or teleology is then *not* something that pertains to the supersensible (as we have to call the object of the intuitive understanding), it rather pertains to our discursivity. So, at least so far, purpose does not apply to the supersensible. Therefore mechanism is not the sole principle applying to the sensible. It has to co-exist with the maxim of final causes in any judgment concerning organized beings (excepting cases where entirely abiological aspects of such beings are in focus). This is why there still remains a question as to the reconcilability of the maxims, in contrast to what would be the case if they were partitioned to separate “respects,” according to the model of the third antinomy.<sup>62</sup>

In a way, the question of reconciliation of the maxims is already solved. As soon as the tendency to mistake them for constitutive principles is exposed, and the transcendental source of their regulativity shown, what remains unclear (apart from the issue of how the maxims are supposed to be used in concrete cases) is only a formal problem: that the maxims are contradictory. This may seem quite serious, but Kant takes the matter lightly. He avoids the contradiction by reinterpreting the maxims. In § 70, he presents the mechanistic maxim as saying that “I *ought* at all times to reflect upon these things *according to the principle* of the simple mechanism of

---

<sup>61</sup> Guyer 2001, 275.

<sup>62</sup> This is clearly stated in Effertz 1994, 252.



nature” (CJ 387).<sup>63</sup> And further, “this does not block the second maxim” according to which I should on occasion, prompted by certain natural forms, reflect on them in terms of final causes (CJ 388). Mechanism is no longer to be judged as the exclusively relevant consideration concerning the possibility of organic beings.

But if the maxims are what they were first said to be, how can they suddenly be changed? And if they can be changed so that no conflict between them arises, why were they not presented correctly from the start? Such considerations can easily lead to the view that the entire antinomy is an artificial construction without real importance. In light of the present reading, Kant’s apparent indifference to the formal contradiction between maxims is rather to be seen as a natural consequence of his way of solving it. To formulate the maxim of mechanism as a proposition claiming an exclusive point of view for judging, as in T, is already an expression of the tendency to take the maxims as objective, constitutive principles. It is only with the help of critique that the recommendation of a mechanistic research strategy can be clearly separated from a commitment to an ontological view regarding how the objects under study were ultimately produced. When such a separation is accomplished, we are free to formulate the maxims as complementary guidelines, rather than as conflicting assertions.<sup>64</sup>

So with the deepened view on the regulative status of the maxims provided in §§ 76–77 their reconciliation now seems to be achieved, and the interpretations of them as compatible offered in § 70 have been given a justification.

But a problematic aspect of the solution is that the function of the idea of a higher understanding may seem somewhat unstable. It has frequently been remarked that notions of a higher understanding are used for rather different purposes in different contexts in the *Critique*

---

<sup>63</sup> Following Allison’s translation of this sentence as well as his treatment of the issue (1991, 30–31).

<sup>64</sup> Compare Allison (1991, 30): “the import of the claim that the principles in question are mere maxims of reflection is that they are to be regarded as guidelines or directives rather than assertions of either ontological or epistemic possibility”.

*of the Power of Judgment*.<sup>65</sup> Whereas § 77 operates with a contrastive *Gegenbild*<sup>66</sup> to our discursivity, aiming to show that what we regard as purposive would not be such from the point of view of an intuitive understanding, Kant at the end of § 77 and in § 78 seems to shift to a more traditional notion of a higher understanding as ultimate ground for the world's purposiveness: "in accordance with the constitution of the human cognitive faculty it is necessary to seek the highest ground of such connections [in terms of purposes] in an original understanding as cause of the world" (CJ 410). Thompson (1995) finds an unresolved ambiguity in the "slippage" from a contrastive intuitive understanding in which purposiveness disappears to one considered as an architect purposively producing the order of the world.

One way to explain this shift is to view it as a question of perspective. Much of the discussion in § 78 concerns the impossibility of reconciling the principles if taken as constitutive (T' and A'), returning to the starting-point of the whole argument. To reconcile them, or at least to reconcile their interpretations as teleology and mechanism understood as ontological principles, we have to posit something that "lies outside of both," the supersensible (CJ 412). But this ground, being entirely indeterminate, does not enable us to explain anything about the possibility of using the two principles; it only serves as an assurance that "at least the possibility that both [principles] may be objectively unifiable in one principle (since they concern appearances that presuppose a supersensible ground) is secured" (CJ 413). This is perhaps not much of a solution: objective unification of the principles *may* be possible on a – for us incomprehensible – supersensible plane. But then the problem to be solved is perhaps not that much of a problem either: it concerns the remaining worry that a conflict may arise between objective principles which are presumed to back up, or follow from, the corresponding regulative principles. And if it is logically possible, for all we know, that the supersensible ground *may* unify teleology and mechanism ontologically, then at least this worry should not hinder us

---

<sup>65</sup> For example by Düsing 1986, 68, Baum 1990, 168–169, Allison 1991, 38–39 and 2000, 87–88, Effertz 1994, 246–247, Thompson 1995, 445–452.

<sup>66</sup> Düsing 1986, 68.

*The Antinomy of Teleological Judgment*

from using the principles regulatively, as maxims of reflective judgment. Now, in CJ 414 it is said that

since the ground of this unifiability lies in that which is neither the one nor the other (neither mechanism nor connection to an end) but is the supersensible substratum of nature, of which we can cognize nothing, the two ways of representing the possibility of such objects are not to be fused into one for our (human) reason, but rather we cannot judge them other than as a connection of final causes grounded in a supreme understanding.

Here we can recognize the two conceptions of a higher understanding that Thompson regards as problematic. The possibility of compatibility is enough to solve the antinomy, thereby swallowing purposiveness in an indeterminable supersensible; but we still have to consider some objects in terms of final ends (conceived on the model of intentional production), since our discursive understanding has no other way of handling the finalistic features of organic beings. The indeterminate supersensible, on the one hand, is conceived from a philosophical perspective, in order to resolve the antinomy. The notion of a supreme understanding as architect of nature, on the other hand, is an idea regulatively used by the power of judgment, for the sake of investigating nature. The first notion belongs to the very abstract perspective of transcendental philosophy, whereas the second one serves as a maxim in the study of biology, taking the form of an assumption to the effect that any structure found in an organism will be found to have some function. Perhaps a conflict between these perspectives might come forth when a Kantian philosopher studies an organism, but presumably it can be appeased by considering that at least from the point of view of transcendental philosophy the functional stance applied in the observation of nature has a merely reflective (non-determinative) status.<sup>67</sup>

One might also wonder what it means in practice to use both maxims, the teleological and the mechanistic, in investigating an organism. I will not pursue the issue in any detail, but only indicate a general point Kant makes about it: that a consideration in terms of

---

<sup>67</sup> As I attempt to show in the preceding essay, functional attributions are regulative from the meta perspective, but from the biologist's object level perspective they are treated as objective.

*The Antinomy of Teleological Judgment*

purposes presupposes some *means* by which the end is possible (CJ 414). As the concept of purpose is that of an object the concept of which is regarded as the cause of the object (CJ 220), there is room for efficient causes as means for realizing the object. For example, if some hitherto unexamined internal structure of an animal body is considered according to final ends, as if it were designed, it is viewed as an organ. This prompts us to search for its function. If a function is hypothetically assumed, we start to investigate the mechanisms by which it operates. In this way a teleological consideration leads us to pursue mechanistic explanation as far as we possibly can (CJ 415). The justification of the regulative concept of a natural purpose is therefore also a justification of the application of mechanism in the life sciences, and this is considered by Kant to be of importance for the scientific status of the enterprise, since without mechanism “no insight into the nature of things can be attained” (CJ 410).

## 7. MECHANISM AND THE INEXPLICABILITY OF ORGANISMS

There are some thorny questions about the status of the principle of mechanism raised by the *Dialectic of the Teleological Power of Judgment*. Some concern the principle itself: What is its status (is it constitutive or regulative)? What is its relation to the causal principle? Closely connected to these is the question of how Kant can claim both that organisms are mechanically unexplainable and that it nonetheless is possible that they are mechanically produced. These views are expressed vividly in the following passage:

For it is quite certain that we can never adequately come to know the organized beings and their internal possibility in accordance with merely mechanical principles of nature, let alone explain them; and indeed this is so certain that we can boldly say that it would be absurd for humans even to make such an attempt or to hope that there may yet arise a Newton who could make comprehensible even the generation [*Erzeugung*] of a blade of grass according to natural laws that no intention has ordered;

*The Antinomy of Teleological Judgment*

rather, we must absolutely deny this insight to human beings. But for us to judge in turn that even if we could penetrate to the principle of nature in the specification of its universal laws known to us there *could* lie hidden no ground sufficient for the possibility of organized beings without the assumption of an intention (that is, in the mere mechanism of nature)<sup>68</sup> underlying their generation would be presumptuous: for how could we know that? (CJ 400)

Let's first take a look at some of Kant's ways of using the terms "mechanism" and "mechanical."<sup>69</sup> In the Introduction to the B-edition of CPR, the "mechanism of nature" is opposed to freedom (CPR Bxxix). In the *Critique of Practical Reason*, also in contrast to freedom, it is said that "all necessity of events in time, according to the natural law of causality, can be called the *mechanism* of nature."<sup>70</sup> The connection of mechanism and causality is very close in this passage, but the contrast to freedom does not quite fit the case of natural purposes, as the solution to the antinomy of teleological judgment is not simply that of the *Third Antinomy*.

In the *Metaphysical Foundations of Natural Science* (1786), Kant criticizes the "mechanical philosophy of nature," according to which the shapes and combinations of atoms or corpuscles, moved by external forces, are to explain the specific variety of matters (*spezifisch verschiedenen Materien*).<sup>71</sup> He prefers a dynamical conception that accords internal forces (attraction and repulsion) to matter. That this sense of "mechanism" is not the one intended in CJ when it is said that organisms cannot be mechanically explained is rather obvious, since in Kant's view such a mechanical philosophy of nature is no good even for physics. A more promising distinction is the one made a few pages earlier in the *Metaphysical Foundations* between mechanical and chemical influence.<sup>72</sup> Mechanical influence is a transmission of movement from one body to another, whereas chemical influence acts also in bodies at rest. One might think that

---

<sup>68</sup> The phrase in parenthesis is missing in Guyer's and Matthews' translation (Kant 2000, 271).

<sup>69</sup> See also the accounts in Allison 1991, 26–27 and Ginsborg 2001, 238–242.

<sup>70</sup> Ak. 5:97.

<sup>71</sup> Ak. 4:532.

<sup>72</sup> Ak. 4:530.

this is the relevant contrast for our case, particularly in view of Kant's well known verdict on chemistry in the preface of the *Metaphysical Foundations*, where chemistry is said to be a systematic art rather than a science, due to its merely empirical regularities lacking *a priori* principles. If chemistry (which for Kant at this time is of a pre-Lavoisierian, poorly mathematized kind) lacks the certainty of physics, perhaps this could be a reason to assign merely reflective status to its explanations? The "Newton of the blade of grass" would then be unable to explain the plant because he works mechanically and not chemically, whereas a chemical explanation (more apt for processes such as growth and reproduction) would count as merely reflectively valid. Against such an interpretation, however, decisive textual evidence from CJ is available. In discussing crystallization, Kant says that the explanation of such formations requires the assumption of no other principle than the mechanism of nature, despite their apparent purposiveness (CJ 348). In this description of crystallization he also states that it concerns true fluids, where the matter is completely dissolved, and he uses the notion of *caloric* (*Wärmestoff*). This clearly shows it to be a chemical process. Such dissolution is exactly what Kant in the *Metaphysical Foundations* calls chemical influence.<sup>73</sup> It must therefore be a broader notion of mechanism, one that includes chemistry but excludes any reference to purposes, that in CJ is said to be unable to account for the origin and development of organisms.<sup>74</sup> As Ginsborg points out, "mechanical" sometimes just means "nonteleological" (as in CJ 406, where "the mechanism of nature" is equated with "a causal connection for which an understanding does not have to be exclusively assumed as a cause").<sup>75</sup>

---

<sup>73</sup> Ak. 4:530.

<sup>74</sup> Compare R4552, Ak. 17:591: "Es ist die Frage, ob es eine organisch bildende Natur gebe (epigenesis) oder bloß eine, die mechanisch und chemisch bildet."

<sup>75</sup> Ginsborg 2001, 239. Ginsborg also says (2001, 239) that even the teleological *Bildungstrieb* (formative drive), postulated by Blumenbach, is an example of "natural mechanism" for Kant, referring to CJ 424. But what Kant actually says in this passage is that the formative drive (as based on a principle of original organization) is distinguished from the "merely mechanical formative power [*Bildungskraft*] that is present in all matter"; the

### *The Antinomy of Teleological Judgment*

This broad notion of mechanism is also easy to find elsewhere in CJ. For instance, in § 77 the causality pertaining to the natural laws of matter is contrasted to causality in terms of purposes (CJ 408), and in § 78 the “mere mechanism of matter” is exemplified by putrefaction (at the time considered to be an exclusively chemical process) (CJ 411). In § 82, the hypothesis that geological revolutions (which certainly involve chemical forces) may have produced organisms is referred to as mechanism (CJ 427–428).

The next question is why mechanism considered as explanatory procedure in CJ is taken to be regulative? One would suppose that, given the causal principle established in the *Second Analogy* of CPR, and the resolution of its *Third Antinomy*, there is no exception to mechanism in nature (freedom being possible only in a perspective different from that of nature). The causal principle is constitutive; how can mechanism then be regulative?

One obvious possibility (alternative (iii) in section 2 above), is that the concepts of causality and mechanism are not equivalent. Given Kant’s flexible use of “mechanism” this would not preclude that he sometimes could take the terms as synonyms, as in the already cited passage from the *Critique of Practical Reason*: “all necessity of events in time, according to the natural law of causality, can be called the *mechanism* of nature.”<sup>76</sup> McLaughlin suggests that there is a development in Kant’s use of “mechanism,” from the first edition of CPR, where the term plays no important role, to its systematic identification with causality in the *Critique of Practical Reason*.<sup>77</sup> But in CJ, written not long after the second *Critique*, McLaughlin finds a new distinction drawn between mechanism and causality, mechanism now being seen as only a particular species of causality in general. The *differentia specifica* of mechanism is that in its explanations the parts determine the whole. Such an “inclusion in space” is a feature that does not follow from the general concept of causality, which only determines a temporal order.<sup>78</sup>

---

formative drive, however, though non-mechanical, leaves considerable room for natural mechanism.

<sup>76</sup> Ak. 5:97.

<sup>77</sup> McLaughlin 1990, 155.

<sup>78</sup> McLaughlin 1990, 152–153.

*The Antinomy of Teleological Judgment*

On McLaughlin's interpretation, it was "the attempt to extend causal explanation in its reductionistic form to the explanation of the organism, that forced Kant to acknowledge that he had ascribed as self-evident a determination to causality that cannot be derived analytically from the concept of cause".<sup>79</sup> So only after equating causality and mechanism in the second *Critique* was Kant in a position to discover that the two concepts, that of a temporally ordered causality and that of explanation of wholes from parts, are not equivalent as he had assumed, and must be distinguished in order to handle the special part-whole relation found in organisms.

A problem for this otherwise plausible view on the evolution of Kant's use of "mechanism" towards a specific position as a sub-type under causality in CJ is that only a little later, in the *Progress* essay from the early 1790s,<sup>80</sup> the concept is, again, clearly equivalent to "causality": *Naturmechanism* is equated with the predetermination of every change by the preceding state.<sup>81</sup> Perhaps a short unpublished piece should not be accorded too much importance, but it appears unlikely that Kant would just ignore his newly reached terminological improvement.

As for the antinomy of teleological judgment, the distinction enables McLaughlin to explain the regulative status of mechanism. Causality is constitutive, but the reductionistic course of explaining wholes from their parts is a human peculiarity, as such merely regulative, and in no way precludes the possibility in principle of holistic explanation of parts in terms of wholes.<sup>82</sup> That reductionism is necessary for us, while holism, though in itself possible (by virtue of being logically compatible with the causal principle), is impossible for us, just is that peculiarity of our understanding, its discursivity, indicated by Kant in § 77. McLaughlin connects this peculiarity to the epistemological situation occasioned by the scientific developments in the 17th century. The methodology of the new science consisted in

---

<sup>79</sup> McLaughlin 1990, 155.

<sup>80</sup> *Welches sind die wirklichen Fortschritte, die die Metaphysik seit Leibnizens und Wolf's Zeiten in Deutschland gemacht hat?* The essay, written for a competition announced by the Royal Academy of Sciences in Berlin but not completed, was published by Rink in 1804.

<sup>81</sup> *Fortschritte*, Ak. 20:289, see also Ak. 20:291.

<sup>82</sup> McLaughlin 1990, 153.



the reductionistic explanation of a phenomenon by the properties and interactions of the parts, with the assumption that *all* properties of a whole should in principle depend on the properties of the parts.<sup>83</sup>

On Kant's view, according to McLaughlin, this "method of classical modern physics is equated with scientific explanation in general and the latter is equated with knowledge as such."<sup>84</sup> It is not quite clear, however, how Kant could claim any necessity (if only for us) in this equation. Furthermore, this reductionistic method is rather similar to the mechanical philosophy of nature criticized by Kant in the *Metaphysical Foundations of Natural Science*. This method, also called *Korpuskularphilosophie*, consists in explanations from "the constitution and composition of [the] smallest parts, as machines,"<sup>85</sup> and it appears to be the same corpuscularianism that McLaughlin, citing Robert Boyle, presents as an example of mechanism.<sup>86</sup> As I have shown above, Kant's mechanism in CJ is more inclusive, adding dynamical and chemical powers.

In § 77, mechanism as explanatory procedure is characterized as going from part to whole and linked to discursivity. McLaughlin construes this as an equation of discursivity with reductionism, and considers the contrasting intuitive understanding as an understanding that would explain in the opposite direction, holistically.<sup>87</sup> While it is difficult to prove that Kant does not here use the terms "discursive" and "intuitive" in a sense restricted to the problem at hand, hence imagining a contrast to our understanding only with respect to part-whole relations, it seems odd that he should have chosen to label such an understanding "intuitive". According to McLaughlin, the intuitive understanding explains holistically, but still in accordance with the causal law (its holism being a species under the genus of causality).<sup>88</sup> But in § 76, the intuitive (*anschauend*) understanding introduced (CJ 402) is clearly a radically different understanding, of the kind discussed in other works by Kant, since this unified cognitive faculty

---

<sup>83</sup> McLaughlin 1990, 174–175.

<sup>84</sup> McLaughlin 1990, 176.

<sup>85</sup> Ak. 4:532–533. A machine is defined as a body or particle whose moving force depends on its figure.

<sup>86</sup> McLaughlin 1990, 176.

<sup>87</sup> McLaughlin 1990, 164–173.

<sup>88</sup> See McLaughlin 1990, 170.

*The Antinomy of Teleological Judgment*

would have no way of distinguishing actuality from possibility. This differs from the intuitive understanding in § 77 which, on McLaughlin's construal, can separate concepts from intuitions, and is dissimilar from us only in that it is able to explain parts from wholes. It is of course possible that another intuitive understanding is discussed in § 77 than the one in § 76,<sup>89</sup> but the fact that Kant connects the peculiarity of our understanding to its discursivity (CJ 406–408) speaks strongly against this suggestion. A discursive understanding thinks by means of concepts.<sup>90</sup> Discursive cognition pertains to conceptual form in abstraction from all content, which amounts to formal logic (CPR A131/B170). To obtain contentful cognition, the discursive understanding depends on data brought forth by an independent receptivity. Consequently, the intuitive understanding, conceived “merely as not discursive” (CJ 406), cannot operate with concepts. Thus the intuitive understanding in § 77 is the same one as that of § 76. Not only mechanism, but also causality (as based on a pure concept of the understanding) must be cancelled for such an understanding.<sup>91</sup>

As discussed above, the broad notion of mechanism employed in the *Antinomy* includes all actions of non-teleological forces of nature. According to Ginsborg, organisms are mechanically inexplicable because the regularities they exhibit cannot be derived from the “lawlikeness displayed by the workings of inorganic nature.”<sup>92</sup> To regulatively regard them as natural purposes amounts to regarding them as normatively constrained. This brings to the otherwise incomprehensible regularities of the organism (what Kant refers to as its contingency) a “lawlike” order.<sup>93</sup> We characterize the organism in terms of biological functions, and this involves a normative view of what its various organs are supposed to do. There is some difficulty on Ginsborg's interpretation as to why only organisms are viewed

---

<sup>89</sup> “Kant's contrast understandings – when they get a name at all – are always called ‘intuitive’ understandings,” says McLaughlin (1990, 170), implying that they have nothing in common but their roles as contrasts.

<sup>90</sup> See for instance CPR A68/B93, and *Prolegomena*, Ak. 4:333.

<sup>91</sup> Compare CPR B312. For a contrary view on the intuitive understanding in § 77, see Peter 1992, 251–258.

<sup>92</sup> Ginsborg 2001, 252.

<sup>93</sup> Ginsborg 2001, 253.

normatively. As she notes, there are also other natural phenomena which exhibit regularities that we are unable to derive from basic laws of physics and chemistry. Perhaps the regularity in the form of a snowflake can be derived from the characteristics of water and air, but many chemical occurrences are more difficult to derive from fundamental laws.<sup>94</sup> Why are we not using teleological conceptions in describing them? Ginsborg's answer is that there is a difference of order between biological and chemical phenomena. To find regularities with regard to the latter, experience is required. That is the case also for the organism, but in addition, its very existence amounts to a regularity that has to be accounted for.<sup>95</sup> Though this distinction does not seem entirely clear, it fits with the tenor of Kant's descriptions in CJ §§ 64–65 of the distinctive self-organizing characteristics of organisms.

Mechanism, then, is regulative because it represents the attempt to reach a complete determination in terms of causality of something which contains regularities that go beyond what follows from the laws prescribed by the understanding. Already in the case of special empirical laws that do not follow from the *Grundsätze*, reflection by means of the regulative principle of the purposiveness of nature for our understanding is needed for systematization. What is added in the case of the organism is so vast a plurality of particular regularities acting in consonance in the same thing, that we not only have to use the general maxim of purposiveness, but also inevitably must consider the thing in teleological terms as a natural purpose, that is, as internally purposive. With respect to a thing conceived according to this concept, namely an organism, the role of mechanism becomes that of a maxim for the continuing investigation of the various causal processes sustaining it. As the organism is considered to be a natural purpose, it is seen as a teleological unit, to which the maxim of mechanism is regulatively applied for the sake of discovery and explanation. This maxim is a regulative extension of the constitutive category of causality, whereas the teleological maxim contains the idea of purpose which belongs to reason; hence the first one is handed over to the power of judgment from the understanding, the second one from reason, as Kant indicates in § 70 (CJ 386).

---

<sup>94</sup> Ginsborg 2001, 245 and 257.

<sup>95</sup> Ginsborg 2001, 257.

Kant's stress on the vast number or near-infinity of the harmonizing special arrangements found in organized beings bears some resemblance to the views of Leibniz and J. N. Tetens. These philosophers, well known to Kant, conceptualized living beings in terms of infinity. According to Leibniz, the organism is a machine whose organization goes all the way down, infinitely, in contradistinction to artificial machines that have parts which themselves are not machine-like but consist of unorganized matter.<sup>96</sup> Tetens differs somewhat from this characterization: he defines organization as an infinitely complex mechanism ("*Die Organisation ist ein unendlich zusammengesetzter Mechanismus*"),<sup>97</sup> so that the difference between organization and mechanism is quantitative rather than qualitative. This difference notwithstanding, both philosophers appeal to infinity in ways that have some affinity to Kant's view. Kant differs from Leibniz in pointing out that the organism is radically dissimilar to a machine because of its capacity for self-reproduction,<sup>98</sup> and at least in CPR he states that it is an empirical question whether one would encounter non-organic parts somewhere in the division of an organized body (A527/B555), and that the notion of a given, infinitely organized object of cognition is incoherent.<sup>99</sup> From Tetens he differs in that he does not consider the distinction between organized beings and mechanically composed bodies to be merely one of degree, not even on the scale of infinity.

Kant points to the contingency of the coming together of all the special laws involved in the functioning of an organism, and this near-infinity of laws is perhaps the reason for his insistence on the merely regulative character of mechanical explanation. Whereas the causal principle is constitutive of objective experience, mechanical explanation, though based on causal laws, is merely regulative when it treats of the infinite. Now, according to the first *Critique*, there is strictly speaking no infinity with regard to the object, but rather an infinite possibility of empirical regress, which is connected to a regulative principle prescribing that for every conditioned its

---

<sup>96</sup> *Monadology* (1714) § 64, in Leibniz 1951, 546.

<sup>97</sup> Tetens 1979 [1777], 475.

<sup>98</sup> Though in the *Opus postumum* he is willing to call it a machine, as for instance in Ak. 21:186.

<sup>99</sup> See McLaughlin 1990, 108–110.

condition should be sought.<sup>100</sup> This principle is presumably expressed also in the regulative principle of mechanism, with its requirement of continuing successive application. In the third *Critique*, the plurality of special laws harmonizing in sustaining the organism presents a complexity which at least with respect to the “peculiarity” of the human vantage point is infinitely too vast for the mechanical mode of explanation to be anything more than just a regulative principle. It is often, and rightly, remarked that Kant distinguishes the constitutive status of causality for objective experience from the merely regulative status of mechanistic *explanation*.<sup>101</sup> The practical impossibility of grasping an infinite chain, indeed its status as an idea of reason, seems to be the rationale for this distinction. To explain something is not just to know that it has some cause, but to track the specific detail of the causal chains involved. Where the causal chains are infinitely (or near-infinitely) complex, mechanistic explanation is merely regulative.

## 8. A FINAL PROBLEM

A difficulty that was mentioned in passing in section 6 above occurs in Kant’s demonstration of the regulativity of the principle of purposiveness in § 77. The employment of a contrasting conception of an intuitive understanding served to show that organisms are not necessarily cognized in the way characteristic for our understanding. The mere conceivability of a non-discursive alternative suffices in Kant’s view to demote the merism of discursivity from any presumed universality. This merism, unavoidable as it is, must therefore be taken as regulative and not as constitutive for the possibility of experiencing the object. A problem with this solution is that it is not entirely clear why the appeal to a higher understanding in this case should lead to the conclusion that a mode of cognition peculiar to the discursive understanding is regulative, whereas in the case of the

---

<sup>100</sup> See A514/B542 on infinity and A509/B537 on the regulative principle. I speak loosely of “infinity” and “near-infinity” in this rather tentative discussion, bypassing the complexities of Kant’s views on infinity.

<sup>101</sup> See for instance Gfeller 1998, 219.

forms of intuition and the categories the peculiarity of discursivity only leads to the conclusion that its objects of cognition should not be considered as objects in themselves. In a nutshell, how can the subjective principles in the first *Critique* be constitutive, considering that they are peculiarities of a discursive understanding?

The problem is potentially damaging for the Kantian architectonic, based as it is on distributing constitutivity to the general principles for the possibility of nature and regulativity to systemic principles that unite minimal experience into wider connections and also apply to special objects with a certain internal complexity (organisms). And this is not an architectonic distinction with a mere ornamental function. If the distinction between constitutive and regulative is untenable, it should mean either that a principle that emerges from the nature of our discursivity is an *a priori* condition for experience, or that such principles are merely regulative. In the case of purposiveness, there have been adherents to both alternatives among commentators.<sup>102</sup> In addition to a constitutive application of teleology in biology, the first alternative, to mention but one among other consequences, brings with it a collapse of the contrast between the regulative idea of an infallibly moral being and the non-ideal human will for which this idea serves as an ideal. The second alternative annuls any claim about an *a priori* structure of experience, thereby transforming Kant's philosophy into a kind of pragmatism.

Whatever one thinks about the intrinsic merits of these consequences, there is no ground for assuming that Kant wished to abrogate the distinction between constitutive and regulative prin-

---

<sup>102</sup> The first alternative was explored by the biologist and philosopher Driesch (1924, 375). The second line is taken in Ernst 1909, 64, and perhaps in Butts 1984, 261. However, I am not aware of any discussion in which the need for such a revision is linked to the dissimilarity between the conclusions drawn by Kant from the appeal to an intuitive understanding in the first and the third *Critiques* (though the dissimilarity as such is noted by McLaughlin (1990, 170–171), who defends the distinction between constitutive and regulative, and takes the dissimilarity as reason for interpreting the notion of intuitive understanding differently in CPR and CJ). Usually the question is prompted by Kant's assignment of merely regulative status to mechanism, which is (unnecessarily) taken to imply that the causal principle itself must be regulative, so that no basis for a difference in kind between teleology and causality remains.

ciples. Rather, if the distinction is threatened by his strategy for demonstrating the regulativity of the maxims, this must be an unwelcome result. Can the distinction be maintained? Assuming that there is a solution to the difficulty (after all, hermeneutics tells us that we should not too lightly attribute an obvious mistake to an author), the following possibilities spring to mind.

1) Maybe the terms “discursivity” and “intuitive understanding” are used very differently in CJ and CPR. As this would avoid the problem, McLaughlin’s claim that alternative understandings in Kant are always fashioned after the requirements of the discussion at hand should perhaps be reconsidered.

2) Even though both the constitutivity of the categories and the regulativity of the principle of purposiveness are explained in similar ways, as being due to the peculiarity of our discursive understanding, there are criteria for distinguishing their status. For example, the categories have schemata, but it is not clear what a schematic time-determination with regard to purposiveness would be.<sup>103</sup> Moreover, biology as a special domain in the empirical study of nature is not ultimately indispensable in the way that general experience is, which is reason enough to distinguish the status of its enabling condition (teleology) from the status of the transcendental principles of experience.

My reasons for not believing 1) have already been given in section 8. It might be thought that the perspectival approach to interpreting Kant adopted in this book should accord well with the notion of a flexible use of terms such as “discursive.” After all, if “thing in itself” is to be taken as a perspective-dependent term (as claimed in the first essay), why shouldn’t there likewise be different uses of “discursivity” on different levels? In answer to that, it must be stressed that discursivity is an ultimate fact on which a great deal of Kant’s system hinges. It is because of discursivity, that is, because of the twofold nature of our cognitive faculty, that we have no cognition of things in themselves (in the term’s transcendental respect). Even though we

---

<sup>103</sup> Compare Driesch 1924, 374. Even more problematic than dissolving the boundary between regulative and constitutive principles is Van De Pitte’s proposal (1975) to assign the status of form of intuition to purposiveness. Should one for instance expect that there is a special kind of mathematical construction in the medium of a pure intuition of purposiveness?

*The Antinomy of Teleological Judgment*

can define other perspectives in which the notion of thing in itself can be differently applied, the possibility of these perspectives depends on their relation to the fundamental fact that our understanding needs material supplied from another source. To relativize this by making the notion of discursivity itself perspective dependent is more likely to disrupt a perspectival interpretation than to further it. Therefore the interpretative strategy described in 2) is preferable. But admittedly, the situation is not quite clear. Further investigation of the compressed thoughts in §§ 76–77 is certainly needed.



## WORKS CITED

### *Kant's works*

References to Kant either indicate volume and page number of the *Akademie* edition, *Kants gesammelte Schriften, herausgegeben von der Deutschen [formerly Preussischen] Akademie der Wissenschaften* (Berlin: de Gruyter 1902–), or only the page number of this edition preceded by an abbreviation (explained in the essays) for the title of the work referred to. The exception is the *Critique of Pure Reason*, for which the standard convention is followed: “A” before a number refers to the pagination of the first edition (1781), “B” refers to the second edition (1787). Sometimes particular English editions are cited as “Kant” followed by year. The editions used are the following:

- Kant 1987, *Critique of Judgment*, translated by Werner Pluhar, Indianapolis: Hackett
- Kant 1992, *Theoretical Philosophy 1755–1770*, translated by David Walford in collaboration with Ralf Meerbote, Cambridge: Cambridge University Press. (I quote from the following works contained in this volume: *The Only Possible Argument in Support of a Demonstration of the Existence of God*; *Attempt to Introduce the Concept of Negative Magnitudes into Philosophy*; *Inquiry Concerning the Distinctness of the Principles of Natural Theology and Morality (Prize Essay)*; *Dreams of a Spirit-Seer Elucidated By Dreams of Metaphysics*; *On the Form and Principles of the Sensible and the Intelligible World (Inaugural Dissertation)*.)
- Kant 1996, *Practical Philosophy*, translated by Mary Gregor, Cambridge: Cambridge University Press. (I quote from: *An Answer to the Question: What Is Enlightenment?*; *On the Wrongfulness of Unauthorized Publication of Books*; *Groundwork of the Metaphysics of Morals*; *Critique of Practical Reason*.)
- Kant 1997a, *Critique of Pure Reason*, translated by Paul Guyer and Allen Wood, Cambridge: Cambridge University Press

*Works Cited*

- Kant 1997b, *Lectures on Metaphysics*, translated by Karl Ameriks and Steve Naragon, Cambridge: Cambridge University Press. (I quote from *Metaphysik K<sub>2</sub>*.)
- Kant 2000, *Critique of the Power of Judgment*, translated by Paul Guyer and Eric Matthews, Cambridge: Cambridge University Press
- Kant 2002, *Theoretical Philosophy After 1781*, edited by Henry E. Allison and Peter Heath, Cambridge: Cambridge University Press. (I quote from: *Prolegomena to any Future Metaphysics that Will Be Able to Come Forward as a Science* (translated by Gary Hatfield); *Metaphysical Foundations of Natural Science* (tr. by Michael Friedman); *On a Discovery whereby any New Critique of Pure Reason Is to Be Made Superfluous by an Older One* (tr. by Henry Allison); *What Progress Has Metaphysics Made in Germany Since the Time of Leibniz and Wolff?* (tr. by Peter Heath).)

*Other works referred to*

- Adams, Robert Merrihew 1997, "Things in Themselves", *Philosophy and Phenomenological Research* 57, 801–825
- Allison, Henry E. 1980, "Kant's Critique of Spinoza", in Richard Kennington (ed.), *The Philosophy of Baruch Spinoza*, Washington, D.C.: Catholic University of America Press, 199–227
- Allison, Henry E. 1983, *Kant's Transcendental Idealism*, New Haven and London: Yale University Press
- Allison, Henry E. 1990, *Kant's Theory of Freedom*, Cambridge: Cambridge University Press
- Allison, Henry E. 1991, "Kant's Antinomy of Teleological Judgment", *Southern Journal of Philosophy* 30, Supplement, 25–42
- Allison, Henry E. 1996, *Idealism and Freedom*, Cambridge: Cambridge University Press
- Allison, Henry E. 2000, "Is the *Critique of Judgment* 'Post-Critical'?", in Sally Sedgwick (ed.), *The Reception of Kant's Critical Philosophy*, Cambridge: Cambridge University Press, 78–92
- Ameriks, Karl 1982, *Kant's Theory of Mind*, Oxford: Clarendon Press

*Works Cited*

- Ameriks, Karl 2001, "Kant and Short Arguments to Humility", in Predrag Cicovacki (ed.), *Kant's Legacy: Essays in Honor of Lewis White Beck*, Rochester: University of Rochester Press, 167–194
- Amundson, Ron 2000, "Against Normal Function", *Studies in History and Philosophy of Biological and Biomedical Sciences* 31, 33–53
- Amundson, Ron and Lauder, George 1994, "Function Without Purpose", *Biology and Philosophy* 9, 443–469
- Anscombe, G. E. M. 1979, "Under a Description", *Noûs* 13, 219–233
- Baertschi, Christian 2004, *Die deutsche metaphysische Kantinterpretation der 1920er Jahre*, Doctoral dissertation, Universität Zürich
- Baublys, Kenneth K. 1975, "Comments on Some Recent Analyses of Functional Statements in Biology", *Philosophy of Science* 42, 469–486
- Baum, Manfred 1990, "Kants Prinzip der Zweckmässigkeit und Hegels Realisierung des Begriffs", in H.-F. Fulda and R.-P. Horstmann (eds.), *Hegel und die 'Kritik der Urteilskraft'*, Stuttgart: Klett-Cotta, 158–173
- Bäumer, Änne 1996, *Geschichte der Biologie*, vol. 3, Frankfurt am Main: Peter Lang
- Beck, Lewis White 1988 (ed.), *Kant Selections*, New York: Macmillan
- Bedau, Mark 1991, "Can Biological Teleology Be Naturalized?", *Journal of Philosophy* 88, 647–655
- Bedau, Mark 1992, "Where's the Good in Teleology?", *Philosophy and Phenomenological Research* 52, 781–806
- Beiser, Frederick 1987, *The Fate of Reason: German Philosophy from Kant to Fichte*, Cambridge, Mass.: Harvard University Press
- Bergmann, Gustav 1968, "Purpose, Function, Scientific Explanation", in May Brodbeck (ed.), *Readings in the Philosophy of the Social Sciences*, New York: Macmillan, 211–223
- Bigelow, John and Pargetter, Robert 1987, "Functions", *Journal of Philosophy* 84, 181–196
- Bird, Graham 1962, *Kant's Theory of Knowledge*, London: Routledge and Kegan Paul
- Bommersheim, Paul 1927, "Der vierfache Sinn der inneren Zweckmässigkeit in Kants Philosophie des Organischen", *Kant-Studien* 32, 290–309

*Works Cited*

- Buchdahl, Gerd 1989, "Realism and Realization in a Kantian Light", in Eva Schaper and Wilhelm Vossenkuhl (eds.), *Reading Kant*, Oxford: Blackwell, 217–249
- Buchdahl, Gerd 1992, *Kant and the Dynamics of Reason*, Oxford: Blackwell
- Buller, David 1998, "Etiological Theories of Function: A Geographical Survey", *Biology and Philosophy* 13, 505–527
- Buller, David, 1999, "Introduction: Natural Teleology", in David Buller (ed.), *Function, Selection, and Design*, Albany: State University of New York, 1–27
- Butts, Robert E. 1984, *Kant and the Double Government Methodology*, Dordrecht: Reidel
- Canguilhem, Georges 1978, *On the Normal and the Pathological* [*Le normal et le pathologique* 1966], Dordrecht: Reidel
- Centi, Beatrice 2001, "Einbildungskraft und Urteilskraft in der *Kritik der Urteilskraft*", in V. Gerhardt, R.-P. Horstmann and R. Schumacher (eds.), *Kant und die Berliner Aufklärung: Akten des IX. internationalen Kant-Kongresses*, vol. 2, Berlin: de Gruyter, 277–282
- Christensen, Wayne D. and Bickhard, Mark H. 2002, "The Process Dynamics of Normative Function", *Monist* 85, 3–28
- Cohen, Gerald Allan 1978, *Karl Marx's Theory of History: A Defence*, Princeton: Princeton University Press
- Cummins, Robert 1975, "Functional Analysis", *Journal of Philosophy* 72, 741–765
- Davies, Paul Sheldon 2000a, "Malfunctions", *Biology and Philosophy* 15, 19–38
- Davies, Paul Sheldon 2000b, "The Nature of Natural Norms: Why Selected Functions Are Systemic Capacity Functions", *Noûs* 34, 85–107
- De Jong, Willem 1995, "Kant's Analytic Judgments and the Traditional Theory of Concepts", *Journal of the History of Philosophy* 33, 613–641
- Deleuze, Gilles 1984, *Kant's Critical Philosophy: The Doctrine of the Faculties* [*La philosophie critique de Kant* 1963], Minneapolis: University of Minnesota Press
- Dennett, Daniel 1993, "Back from the Drawing Board", in Bo Dahlbom (ed.), *Dennett and His Critics*, Oxford: Blackwell, 203–235

*Works Cited*

- Dörflinger, Bernd 1995, "The Underlying Teleology of the First Critique", in H. Robinson (ed.), *Proceedings of the Eighth International Kant Congress*, vol. I, Milwaukee: Marquette University Press, 813–826
- Driesch, Hans 1924, "Kant und das Ganze", *Kant-Studien* 29, 365–376
- Düsing, Klaus 1986, *Die Teleologie in Kants Weltbegriff*, second edition, Bonn: Bouvier
- Effertz, Dirk 1994, *Kants Metaphysik: Welt und Freiheit*, Freiburg/München: Verlag Karl Alber
- Erdmann, Benno 1878, *Kants Kriticismus in der ersten und in der zweiten Auflage der Kritik der reinen Vernunft*, Leipzig: Verlag von Leopold Voss
- Ernst, Wilhelm 1909, *Der Zweckbegriff bei Kant und sein Verhältnis zu den Kategorien*, Berlin: Reuther und Reichard
- Falkenstein, Lorne 1995, *Kant's Intuitionism. A Commentary on the Transcendental Aesthetic*, Toronto: University of Toronto Press
- Frost, Walter 1906, "Kants Teleologie", *Kant-Studien* 11, 297–347
- Gawlina, Manfred 1996, *Das Medusenhaupt der Kritik*, Berlin: de Gruyter
- Gfeller, Thomas 1998, "Wie tragfähig ist der teleologischen Brückenschlag?", *Zeitschrift für philosophische Forschung* 52, 215–236
- Ginsborg, Hannah 2001, "Kant on Understanding Organisms as Natural Purposes", in Eric Watkins (ed.), *Kant and the Sciences*, Oxford: Oxford University Press, 231–258
- Godfrey-Smith, Peter 1994, "A Modern History Theory of Functions", *Noûs* 28, 344–362
- Grier, Michelle 2001, *Kant's Doctrine of Transcendental Illusion*, Cambridge: Cambridge University Press
- Griffiths, Paul E. 2001, "Genetic Information: A Metaphor in Search of a Theory", *Philosophy of Science* 68, 394–412
- Guyer, Paul 2001, "Organisms and the Unity of Science", in Eric Watkins (ed.), *Kant and the Sciences*, Oxford: Oxford University Press, 259–281
- Haffner, Thomas 1997, *Die Epigenesisanalogie in Kants Kritik der reinen Vernunft*, Doctoral dissertation, Universität des Saarlandes
- Heidegger, Martin 1991, *Kant und das Problem der Metaphysik*, [first edition 1929], Frankfurt am Main: Vittorio Klostermann

*Works Cited*

- Heidegger, Martin 1995, *The Fundamental Concepts of Metaphysics* [*Die Grundbegriffe der Metaphysik* 1983], Bloomington: Indiana University Press
- Henrich, Dieter 1989, “Kant’s Notion of a Deduction and the Methodological Background of the First *Critique*”, in E. Förster (ed.), *Kant’s Transcendental Deductions. The Three Critiques and the Opus postumum*, Stanford: Stanford University Press, 29–46
- Heyse, Hans 1927, *Der Begriff der Ganzheit und die kantische Philosophie*, München: Verlag von Ernst Reinhardt
- Hinske, Norbert 1965, “Kants Begriff der Antinomie und die Etappen seiner Ausarbeitung”, *Kant-Studien* 56, 485–496
- Högemann, Brigitte 1980, *Die Idee der Freiheit und das Subjekt*, Königstein/Ts.: Hain
- Ingensiep, Hans Werner 1994, “Die biologischen Analogien und die erkenntnistheoretischen Alternativen in Kants *Kritik der reinen Vernunft* B § 27”, *Kant-Studien* 85, 381–393
- Jacob, François 1993, *The Logic of Life. A History of Heredity*, Princeton: Princeton University Press
- Jardine, Nick 2003, “Hermeneutic Strategies in Gerd Buchdahl’s Kantian Philosophy of Science”, *Studies in History and Philosophy of Science* 34, 183–208
- Jolley, Nicholas 1988, “Leibniz and Malebranche on Innate Ideas”, *Philosophical Review* 97, 71–91
- Kemp Smith, Norman 2003, *A Commentary to Kant’s Critique of Pure Reason* [reprint of the second edition 1923], Basingstoke: Palgrave Macmillan
- Kirschner, Marc, Gerhart, John and Mitchison, Tim 2000, “Molecular ‘Vitalism’”, *Cell* 100, 79–88
- Kleingeld, Pauline 1998, “The Conative Character of Reason in Kant’s Philosophy”, *Journal of the History of Philosophy* 36, 77–97
- Klemme, Heiner F. 1996, *Kants Philosophie des Subjekts*, Hamburg: Felix Meiner Verlag
- Kripke, Saul 2001, “Identity and Necessity”, in Michael J. Loux (ed.), *Metaphysics: Contemporary Readings*, London: Routledge, 218–247
- Kroon, Frederick 2001, “The Semantics of ‘Things in Themselves’: A Deflationary Account”, *Philosophical Quarterly* 51, 165–181

*Works Cited*

- Kuehn, Manfred 1997, "The Wolffian Background of Kant's Transcendental Deduction", in Patricia Easton (ed.), *Logic and the Workings of the Mind*, Atascadero: Ridgeview, 229–250
- Langsam, Harold 2000, "Kant's Compatibilism and His Two Conceptions of Truth", *Pacific Philosophical Quarterly* 81, 164–188
- Langton, Rae 1998, *Kantian Humility. Our Ignorance of Things in Themselves*, Oxford: Clarendon Press
- Leibniz, Gottfried Wilhelm 1951, *Selections*, Philip. P Wiener (ed.), New York: Charles Scribner's Sons
- Lennox, James 2001, *Aristotle's Philosophy of Biology*, Cambridge: Cambridge University Press
- Leppäkoski, Markku 1993, *The Transcendental How. Kant's Transcendental Deduction of Objective Cognition*, Stockholm: Almqvist & Wiksell International
- Louguenesse, Béatrice 2000, "Point of View of Man or Knowledge of God: Kant and Hegel on Concept, Judgment, and Reason", in Sally Sedgwick (ed.), *The Reception of Kant's Critical Philosophy*, Cambridge: Cambridge University Press, 253–282
- Lovejoy, Arthur O. 1911, "Kant and Evolution II", *The Popular Science Monthly* 78, 36–51
- McCarthy, Michael H. 1985, "The Objection of Circularity in Groundwork III", *Kant-Studien* 76, 28–42
- McFarland, John D. 1970, *Kant's Concept of Teleology*, Edinburgh: University of Edinburgh Press
- Machamer, Peter 1977, "Teleology and Selective Processes" in Robert G. Colodny (ed.), *Logic, Laws, and Life*, Pittsburgh: University of Pittsburgh Press, 129–142
- McLaughlin, Peter 1989, "What Is an Antinomy of Judgment?", in G. Funke and Th. M. Seebohm (eds.), *Proceedings: Sixth International Kant Congress*, vol. II/2, Washington, D.C.: University Press of America, 357–367
- McLaughlin, Peter 1990, *Kant's Critique of Teleology in Biological Explanation: Antinomy and Teleology*, Lewiston: The Edwin Mellen Press
- McLaughlin, Peter 2001, *What Functions Explain. Functional Explanation and Self-Reproducing Systems*, Cambridge: Cambridge University Press

*Works Cited*

- Makkreel, Rudolf 1990, *Imagination and Interpretation in Kant*, Chicago: University of Chicago Press
- Makkreel, Rudolf 1991, "Regulative and Reflective Uses of Purposiveness in Kant", *Southern Journal of Philosophy* 30, Supplement, 49–63
- Manning, Richard 1997, "Biological Function, Selection, and Reduction", *British Journal for the Philosophy of Science* 48, 69–82
- Marc-Wogau, Konrad 1938, *Vier Studien zu Kants Kritik der Urteilskraft*, Uppsala: A.-B. Lundequistska Bokhandeln
- Matthews, H. E. 1969, "Strawson on Transcendental Idealism", *Philosophical Quarterly* 19, 204–220
- Maund, Barry 2000, "Proper Functions and Aristotelian Functions in Biology", *Studies in History and Philosophy of Biological and Biomedical Sciences* 31, 155–178
- Maynard Smith, John 2000, "The Concept of Information in Biology", *Philosophy of Science* 67, 177–194
- Meerbote, Ralf 1974, "The Unknowability of Things in Themselves", in Lewis White Beck (ed.), *Kant's Theory of Knowledge*, Dordrecht: Reidel, 166–174
- Melander, Peter 1997, *Analyzing Functions*, Stockholm: Almqvist & Wiksell International
- Millikan, Ruth Garrett 1989a, "Biosemantics", *Journal of Philosophy* 86, 281–297
- Millikan, Ruth Garrett 1989b, "In Defense of Proper Functions", *Philosophy of Science* 56, 288–302
- Melnick, Arthur 1973, *Kant's Analogies of Experience*, Chicago: Chicago University Press
- Milz, Bernhard 2002, *Der gesuchte Widerstreit. Die Antinomie in Kants 'Kritik der praktischen Vernunft'*, Berlin: de Gruyter
- Mitchell, Sandra 1993, "Dispositions or Etiologies? A Comment on Bigelow and Pargetter", *Journal of Philosophy* 90, 249–259
- Mitchell, Sandra 1995, "Function, Fitness and Disposition", *Biology and Philosophy* 10, 39–54
- Moss, Lenny 1992, "A Kernel of Truth? On the Reality of the Genetic Program", *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, vol. 1: Contributed Papers, 335–348



*Works Cited*

- Nagel, Gordon 1983, *The Structure of Experience*, Chicago: University of Chicago Press
- Neander, Karen 1991a, "Functions as Selected Effects: The Conceptual Analyst's Defense", *Philosophy of Science* 58, 168–184
- Neander, Karen 1991b, "The Teleological Notion of 'Function'", *Australasian Journal of Philosophy* 69, 454–468
- Neander, Karen 2002, "Types of Traits: The Importance of Functional Homologues", in André Ariew, Robert Cummins and Mark Perlman (eds.), *Functions. New Essays in the Philosophy of Psychology and Biology*, Oxford: Oxford University Press, 390–415
- Neon, Thomas 1986, *Objektivität und endliche Erkenntnis. Kants transzendentalphilosophische Korrespondenztheorie der Wahrheit*, Freiburg/München: Verlag Karl Alber
- Nissen, Lowell 1993, "Four Ways of Eliminating Mind from Teleology", *Studies in History and Philosophy of Science* 24, 27–48
- Oberhausen, Michael 1997, *Das neue Apriori: Kants Lehre von einer 'ursprünglichen Erwerbung' apriorischer Vorstellungen*, Stuttgart/Bad Cannstatt: Frommann-Holzboog
- O'Neill, Onora 1989, *Constructions of Reason*, Cambridge: Cambridge University Press
- O'Neill, Onora 1992, "Vindicating Reason", in Paul Guyer (ed.), *The Cambridge Companion to Kant*, Cambridge: Cambridge University Press, 280–308
- O'Shea, James 1999, Review of Langton 1998, *International Journal of Philosophical Studies* 7, 253–257
- Palmquist, Stephen 1993, *Kant's System of Perspectives*, Lanham: University Press of America
- Palti, Elias 1999, "The 'Metaphor of Life': Herder's Philosophy of History and Uneven Developments in Late Eighteenth-Century Natural Sciences", *History & Theory* 38, 322–347
- Parkinson, G. H. R. 1995, "Philosophy and Logic", in Nicholas Jolley (ed.), *The Cambridge Companion to Leibniz*, Cambridge: Cambridge University Press, 199–223
- Paton, Herbert James 1936, *Kant's Metaphysic of Experience*, vol. I, London: George Allen & Unwin
- Paton, Herbert James 1991, *The Moral Law. Kant's Groundwork of the Metaphysic of Morals*, London: Routledge

*Works Cited*

- Patt, Walter 1987, *Transzendentaler Idealismus: Kants Lehre von der Subjektivität der Anschauung in der Dissertation von 1770 und in der Kritik der reinen Vernunft*, Berlin: de Gruyter
- Peter, Joachim 1992, *Das transzendente Prinzip der Urteilskraft*, Berlin: de Gruyter
- Petzäll, Åke 1933, *Der Apriorismus Kants und die "Philosophia pigrorum"*, Göteborg: Wettergren & Kerbers förlag
- Phalén, Adolf 1976, *Ur efterlämnade manuskript 7. Kants kunskapslära*, Stockholm: Natur och Kultur
- Piché, Claude (forthcoming), "Kant and the Problem of Affection", in P. Fairfield (ed.), *Festschrift for Gary B. Madison*, Ottawa: University of Ottawa Press
- Pistorius, Hermann Andreas 2000, "Elucidations of Professor Kant's Critique of Pure Reason, by Johann Schultze" ["Erläuterungen über des Herrn Professor Kant Kritik der reinen Vernunft von Johann Schultze", *Allgemeine deutsche Bibliothek* 1786], in Sassen 2000, 93–105
- Prauss, Gerold 1977, *Kant und das Problem der Dinge an sich*, Bonn: Bouvier
- Rang, Bernhard 1993, "Zweckmässigkeit, Zweckursächlichkeit und Ganzheitlichkeit in der organischen Natur", *Philosophisches Jahrbuch* 100, 39–71
- Ratcliffe, Matthew 2000, "The Function of Function", *Studies in History and Philosophy of Biological and Biomedical Sciences* 31, 113–133
- Ratcliffe, Matthew 2001, "A Kantian Stance on the Intentional Stance", *Biology and Philosophy* 16, 29–52
- Rescher, Nicholas 2000, *Kant and the Reach of Reason*, Cambridge: Cambridge University Press
- Richards, Robert J. 2000, "Kant and Blumenbach on the *Bildungstrieb*: A Historical Misunderstanding", *Studies in History and Philosophy of Biology and Biomedical Sciences* 31, 11–32
- Robinson, Hoke 1994, "Two Perspectives on Kant's Appearances and Things in Themselves", *Journal of the History of Philosophy* 32, 411–441
- Robinson, Hoke 2001, "The Disappearing X", in V. Gerhardt, R.-P. Horstmann and R. Schumacher (eds.), *Kant und die Berliner Aufklärung: Akten des IX. internationalen Kant-Kongresses*, vol. 2, Berlin: de Gruyter, 418–428

*Works Cited*

- Rosales, Alberto 1989, "Zur teleologischen Grundlage der transzendenten Deduktion der Kategorien", *Kant-Studien* 80, 377–404
- Rosefeldt, Tobias 2001, Review of Langton 1998, *European Journal of Philosophy* 9, 263–268
- Rosenberg, Alexander 1994, *Instrumental Biology or the Disunity of Science*, Chicago: University of Chicago Press
- Ruse, Michael 1973, *The Philosophy of Biology*, London: Hutchinson
- Sassen, Brigitte 2000 (ed.), *Kant's Early Critics*, Cambridge: Cambridge University Press
- Schelling, F. W. J. 1907, *Vom Ich als Prinzip der Philosophie* [1795], in *Werke*, vol. 1, Leipzig: Fritz Eckardt Verlag
- Schlosser, Gerhard 1998, "Self-re-production and Functionality: A Systems-Theoretical Approach to Teleological Explanation", *Synthese* 116, 303–354
- Schönecker, Dieter 1997, "Die 'Art von Zirkel' im dritten Abschnitt der *Grundlegung*", *Allgemeine Zeitschrift für Philosophie* 22, 189–202
- Schönecker, Dieter 1999, *Kant: Grundlegung III. Die Deduktion des kategorischen Imperativs*, Freiburg/München: Verlag Karl Alber
- Schurz, Gerhard 2001, "What is 'Normal'? An Evolution-Theoretic Foundation for Normic Laws and their Relation to Statistical Normality", *Philosophy of Science* 68, 476–497
- Setiya, Kieran 2004, "Transcendental Idealism in the *Aesthetic*", *Philosophy and Phenomenological Research* 68, 63–88
- Siewert, Donald J. 1974, "Kant's Dialectic of Teleological Judgment", in G. Funke (ed.), *Akten des 4. internationalen Kant-Kongresses*, vol. 2:1, Berlin: de Gruyter, 452–460
- Sloan, Phillip R. 2002, "Preforming the Categories: Eighteenth-Century Generation Theory and the Biological Roots of Kant's A Priori", *Journal of the History of Philosophy* 40, 229–253
- Sober, Elliott 1993, *Philosophy of Biology*, Oxford: Oxford University Press
- Sorabji, Richard 1964, "Function", *Philosophical Quarterly* 14, 289–302
- Spindler, Josef 1925, "Zur Frage der Interpretation einer der wichtigsten Stellen der 'Kritik der Urteilskraft'", *Kant-Studien* 30, 468–470

### Works Cited

- Steigleder, Klaus 2002, *Kants Moralphilosophie. Die Selbstbezüglichkeit reiner praktischer Vernunft*, Stuttgart: Verlag J.B. Metzler
- Strawson, P. F. 1966, *The Bounds of Sense*, London: Methuen
- Strawson, P. F. 1994, "The Problem of Realism and the A Priori", in Paolo Parrini (ed.), *Kant and Contemporary Epistemology*, Dordrecht: Kluwer, 167–174
- Tarbet, David W. 1968, "The Fabric of Metaphor in Kant's *Critique of Pure Reason*", *Journal of the History of Philosophy* 6, 257–270
- Tetens, Johann Nicolaus 1779, *Philosophische Versuche über die menschliche Natur und ihre Entwicklung* [reprint of the first edition 1777], vol. II, Hildesheim: Georg Olms Verlag
- Thompson, Kevin 1995, "The Antinomy of Teleological Judgment and the Concept of an Intuitive Intellect: Transformation and Conflict", in H. Robinson (ed.), *Proceedings of the Eighth International Kant Congress*, vol. II, Milwaukee: Marquette University Press, 445–452
- Unzer, Johann August 1782, "Johann Friedrich Blumenbach der Med. Prof. ord. zu Göttingen, über den Bildungstrieb und das Zeugungsgeschäft", *Allgemeine deutsche Bibliothek* 50, 164–173, available at <http://www.ub.uni-bielefeld.de/diglib/aufklaerung/>
- Van Cleve, James 1999, *Problems from Kant*, New York and Oxford: Oxford University Press
- Van Cleve, James 2003, "Précis of *Problems from Kant*", *Philosophy and Phenomenological Research* 66, 190–195
- Van De Pitte, Frederick 1975, "The Role of Teleology in Kant's Work", in W. Werkmeister (ed.), *Reflections on Kant's Philosophy*, Gainesville: University Press of Florida, 135–147
- Wachbroit, Robert 1994, "Normality as a Biological Concept", *Philosophy of Science* 61, 579–591
- Walker, Ralph 2002, Review of Langton 1998, *Mind* 111, 136–143
- Walker, Steven M., Downes, C. Peter and Leslie, Nick R. 2001, "TPIP: A Novel Phosphoinositide 3-phosphatase", *Biochemical Journal* 360, 277–283
- Walsh, D. M. 1996, "Fitness and Function", *British Journal for the Philosophy of Science* 47, 553–574
- Weldon, T. D. 1958, *Kant's Critique of Pure Reason*, second edition, Oxford: Clarendon Press

*Works Cited*

- Westphal, Kenneth R. 1997, "Noumenal Causality Reconsidered: Affection, Agency, and Meaning in Kant", *Canadian Journal of Philosophy* 27, 209–245
- Westphal, Kenneth R. 1998, "Buchdahl's 'Phenomenological' View of Kant: A Critique", *Kant-Studien* 89, 335–352
- Williams, Mary B. 1976, "The Logical Structure of Functional Explanations in Biology", *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, vol. 1: Contributed Papers, 37–46
- Wittgenstein, Ludwig 1977, *Philosophische Untersuchungen*, Frankfurt am Main: Suhrkamp
- Wood, Allen W. 1999, *Kant's Ethical Thought*, Cambridge: Cambridge University Press
- Woodfield, Andrew 1976, *Teleology*, Cambridge: Cambridge University Press
- Wouters, Arno Gerhard 1999, *Explanation Without a Cause*, Utrecht: Zeno
- Wright, Larry 1973, "Functions", *Philosophical Review* 82, 139–168
- Yamane, Yûichirô 2001, "Von der 'Erwerbung' zur 'ursprünglichen Erwerbung' – Ein Querschnitt durch die Entstehungsgeschichte der kritischen Metaphysik", in V. Gerhardt, R.-P. Horstmann and R. Schumacher (eds.), *Kant und die Berliner Aufklärung: Akten des IX. internationalen Kant-Kongresses*, vol. 2, Berlin: de Gruyter, 691–699
- Zammito, John 1992, *The Genesis of Kant's Critique of Judgment*, Chicago: University of Chicago Press
- Zammito, John 2003, "'This Inscrutable Principle of an Original Organization': Epigenesis and 'Looseness of Fit' in Kant's Philosophy of Science", *Studies in History and Philosophy of Science* 34, 73–109
- Zanetti, Véronique 1993, "Die Antinomie der teleologischen Urteilskraft", *Kant-Studien* 83, 341–355
- Zöllner, Günter 1988, "Kant on the Generation of Metaphysical Knowledge", in Hariolf Oberer and Gerhard Seel (eds.), *Kant. Analysen – Probleme – Kritik*, Würzburg: Königshausen & Neumann, 71–90
- Zumbach, Clark 1984, *The Transcendent Science: Kant's Conception of Biological Methodology*, The Hague: Martinus Nijhoff