

Decaan  
Rector

Prof. dr. Marc Boone  
Prof. dr. Anne De Paepe





Faculteit Letteren & Wijsbegeerte

Boris Demarest

# *Kant's Transcendental Organics*

*Systematicity and its Historicity*

Proefschrift voorgelegd tot het behalen van de graad van  
Doctor in de Wijsbegeerte

2016



# Acknowledgements

Over the course of the years I have been engaged in this project, I have accrued so massive an intellectual and personal debt that I cannot but despair giving credit wherever credit is due. I shall follow my intellectual hero Immanuel Kant, however, in his conviction that the unrealizability of an ideal does not absolve one of acting in accordance with it, and accept that many will find this attempt at acknowledgement inadequate. Like Kant, also, I will attempt to attain completeness through a systematization, a division into categories of people to whom I feel indebted. The list such an approach generates does not, as will be clear to any reader, follow the order of a *degree* or *amount* of debt, for that is in most cases impossible to measure and to compare, and in some cases immeasurable and incomparable.

The first person to thank is, of course, my supervisor, Gertrudis Van de Vijver, for her continuing faith in my project and my capacity to bring it to fruition, and her continuing support even in times when my academic future and present were unclear. I have learned more from her than the surface of this text can reveal, and than our possible disagreements may show. I hope therefore that I have in some way met up to her expectations of me and my work, and that I have succeeded in repaying her by learning her a few things myself.

Besides in my supervisor I have always found a great deal of support and intellectual nourishment at the Centre for Critical Philosophy. I owe many of my colleagues there considerably, not just for what they taught me, but also for the constructive and generous way in which they forced me to submit my blossoming conceptions and my manner of expression to careful scrutiny. I feel that this crucial dialogical dimension typical of intellectual encounters at the Centre has found its way in the general style and method of this text. My particular gratitude goes to Emiliano Acosta and Joris Van Poucke. Both of these men has, in his own way, functioned as a mentor of sorts in stages of my development, and I feel grateful for their ultimate willingness to let go of the asymmetry that once characterized our relationships.

It was unfortunately late that I found a way of communicating and exchanging intellectually with members of other research units at the department of Philosophy at

Ghent University. Nevertheless, in the relatively short period that I have had the occasion to work with them in various contexts, these people have left an indelible impression on my thought, my intellectual practice and my moral character. They have also kindly allowed, encouraged and helped me to pursue research that was necessary to this project, but hard to complete given my own training. Of these many intellectual friends, I cannot but specifically thank Laura Georgescu, Barnaby Hutchins and Violi Sahaj for their particular generosity in friendship towards me.

Perhaps special thanks is also due to Eric Schliesser and Charles Wolfe. For several years, they managed to make Ghent an exciting place to do graduate research by frequently organizing workshops, conferences and lectures with scholars from many countries and even more philosophical inclinations. Simply being able to attend such events with such frequency is a continuous source of knowledge and intellectual challenge.

One of the most beneficial aspects of my period of research was, however, that besides my dealings with academic peers and superiors, I had the opportunity to engage with students in the context of various classes. We are tempted to underestimate the value for the development of an idea of having to put it clearly, or of having to get clear on the informational basis on which it is constructed, in order to communicate it to people who do not have, or feel no need to feign, the necessary background.

But it is not just within the once again increasingly fortified walls of academia that I have found intellectual and personal support. I have been blessed with a circle of close friends who have miraculously encouraged me in pursuing this enterprise which should otherwise be alien to them. Given the fact that they are hardly in the position to appreciate the inner workings and dealings of philosophy as a discipline, they have shown remarkable understanding for my passion for and faith in it. As a society, we underestimate the near indispensability of expectations of those close to us to our determination in ambitious enterprises. Of these people, I would like to explicitly thank Pauline Groen, Sebastien Helmoortel, Maarten Puype, Kevin Schoonooghe en Jonas Van Braeckeveldt, for reasons known to them.

Finally, I need to thank my family, starting with those that irreparable time has swept away from me. This project has been instrumental in my process of mourning for my grandparents. Each of the three grandparents that lived to help (or meddle) in raising me, Hendrik Demarest, Cecilia Hillewaert and Eugenie Dewinter, has been more important in my development as an aficionado of intellectual and cultural activity than I could realize when they were still alive. They have done much to supply the child that I once was with means for further development. I can imagine no better way of honouring their memory than by trying to prove them right in their faith in me, and by not allowing their loving efforts to have been in vain.

I have not walked the arduous path of maturation alone, of course, nor do I find myself bereft of all family now. That is why I feel the need to express my gratitude to my

sisters, Leila and Senta Demarest, here. They too have shown much faith in me, and offered me more support at times of distress than my delusions of autonomy allow me to recall. It is in all likelihood what we share that has allowed us to bring up the mutual understanding that is often so hard to come by. Moreover, it is not unimportant that my sister Leila has chosen a similar lifestyle to my own, which offered occasions for us to exchange on the joys and pains of academic labour.

Above all, however, I need to express my massive indebtedness to my parents, Peter Demarest and Elisabeth Janssens. I have the firm conviction that, in bringing a child onto this world, we charge ourselves with an enormous asymmetrical duty towards it – asymmetrical because we do not also charge that child with duties towards us. What is traditionally taken as the “respect” children owe their parents, is no more than a pragmatic duty: in teaching a child, in bringing it up, we need some authority over it in order to cultivate in it the sense of internal limitation that is essential to all autonomy. We can never lay claim, however, to duties of that child towards us beyond these: it does not build up debt to us by being in our charge: we incur a debt when we fail to fulfil our parental duties, and we absolve ourselves of duty only by allowing the child to be an autonomous individual. In light of my endorsement of this radical doctrine, the reader will undoubtedly realize just how far the effort and support of my parents, both emotional and financial, must have reached in order for me to actually believe I have incurred a massive debt towards them. Because of the magnitude of my gratefulness, I can find no better way of expressing it than this.





## Preface

Dans les sciences qui sont difficiles par elles-mêmes, je ne mesure pas la longueur d'un Livre par le nombre de ses pages, mais par la longueur du tems qu'il faut employer pour l'entendre. En ce sens il est assez souvent arrive que l'ouvrage rendu un peu plus long, auroit été beaucoup plus court. (Terrasson 1754: 146)

This saying of Jean Terrasson is famously invoked by Kant in the preface to the first edition of his magnum opus, the *Critique of Pure Reason*. In spite of the fact that Kant goes on to adapt it into a saying of his own, he seems to always have been aware of it, and to have realized just to what extent it was applicable to the work he finally laid before the public after his so-called “silent decade”, for the *Critique of Pure Reason* would have been a much shorter book, had it not been quite so short.

It seems odd to call the *Critique of Pure Reason* a short book, given that it numbered almost 900 pages in its first edition version. Nevertheless, it is safe to say that few of these 900 pages contain no sentence or paragraph on which readers would like Kant to have expanded a bit. The desire to keep the length of the book manageable may thus have ended up thwarting readers' efforts to understand it.

I would like to think I have learned from this error of judgment of Kant's – which is why I have ended up writing a dissertation that is admittedly somewhat too lengthy. The choice was deliberate, for I feared that a concise treatment of the issue would have limited the ease of reading and of appreciating the point I sought to make. The usual style of contemporary philosophy, which is particularly adapted to short journal articles, and monographs consisting essentially of a collage of such articles, would not have been appropriate for my present argument. In this preface, I would like to offer a defense of form, style and method of this dissertation.

One of the reasons why this dissertation has ended up as long as it is, is the constant effort at historical situation. I expect this will be puzzling to some, since the work is not recognizably an effort in antiquarian history of philosophy. I even expect many to charge it with being blatantly anachronistic in the views it ascribes to Kant. This peculiarity is due to the fact that I have constantly sought to historicize in order to de-historicize.

The maxim to read a work historically, and not anachronistically, postulates that we have ready access to the historic context of a work to use as a background for our interpretative practice. The problem with many historicizations is that it is naïve precisely in its picture of such a stable background of interpretation, for an obvious hermeneutic circle occurs: the historical context of the work is usually regarded as the body of texts, the tradition, in which it occurs. But each work in this tradition itself belongs to that body of work. Consequently, a revision in our picture of a work in our tradition implies a revision in our picture of that tradition, which in turn prompts a revision of the individual works composing that tradition. The antiquarian belief consists in the insistence that such holistic reframing is strongly constrained by the historical data, and that historians usually have a solid grasp of the background of a text. I find this belief naïve, especially in the case of a figure such as Kant.

The historic background against which Kant is usually historicized, I submit, is itself highly anachronistic, in two ways. First of all, in our antiquarianism, we take Kant to be responding to what we have made of his contemporaries, not to what he himself made of their texts and views. Second, we read his arguments against the background of his successors, thereby tacitly agreeing with their interpretations of the problems he was dealing with, and with their identification of the merits and demerits of the critical philosophy. In this way, we are reading Kant against our historical reconstruction of his background, not against a stable background. The hermeneutic effort implies that we may have to revise our image of the background in order to understand Kant's place in it.

In this dissertation, I attempt to present the history of debates in which Kant subsequently intervened in the form in which I believe Kant viewed them. In other words, I do not attempt to reconstruct an objective background, but rather a background against which Kant put himself. I usually argue that what contemporary historians of philosophy, of ideas and of science take this background to be makes no sense of Kant's place in it. For example, I try to show in chapter 1 that the challenge of understanding Kant's position towards mechanicism is the result of reading a later, positivistic evaluation of Newton into Kant. In chapter 2, I blame the difficulties in interpreting Kantian epigenesis partly on the false assumption that Kant believed there to be only two options: epigenesis and preformation. This is again a legacy from the late 19<sup>th</sup> century and its debatable, because incontrovertibly ideological, image of its scientific past. In chapter 3, I suggest that there has been a rather complex and heinous dialectic between the interpretation of early modern philosophy through the lens of the empiricism and rationalism-opposition. The dialectic is complex because we usually point to Kant for this image of early modern epistemology. Here, a problematic interpretation of Kant feeds back into a problematic interpretation of the debate in which he was intervening. In chapter 4, I argue that the later emphasis on the *banishment* of teleology from mechanicism is disingenuous, and that Kant was in all

likelihood not fooled by it. Aristotelianism and Mechanicism do not differ in that they take recourse to teleology somehow; they differ in their interpretation and integration of teleology and teleological explanation. In chapter 5, I suggest that we have read Kant's comments on systematicity as if Leibnizianism was obvious to him, and as if he had not read the major challenges to systematicity in the Enlightenment, even though he was at least aware of those offered by Locke and Buffon. As a result, we have overemphasized his recuperation of rationalism, and underappreciated his concessions to what he regarded as empiricism. In chapter 6, I suggest Kant's attack on rational psychology was motivated by a desire to maintain the disunity of the faculties, or rather change the image of how they unite and relate. This contrasts with readings that want to read him as ultimately seeking the unity of the faculties.

In all these reconstructions, one might find inaccuracies and inadequacies. But although I am certainly open to debating them and learning from others regarding these issues, I believe their ultimate adequacy matters less than the light they shed on the problems Kant set himself and how he went about in solving them. It also opens the route to showing that his answers are more original, and less irrevocably time-bound, than is sometimes assumed. For this is the central sin of historicization: it reads our philosophical predecessors as if they were trying to solve issues about which we now no longer care, of which we have been happily rid by the subsequent progress of human knowledge. The problem, thus, is that many so-called antiquarian, historicizing readings evade wiggishness, i.e. reading our past in the light of the present, only by explaining what Hasok Chang (2009) has chastised as "triumphalism", i.e. reading the past as rightfully overcome by subsequent evolutions. And we are rarely more triumphalist than in the history of philosophy.

The triumphalism in Kant scholarship consists in the second anachronistic feature I identified above, namely that of reading his arguments against the background of his successors. This background stems mostly from three sources, all of which are importantly entangled, namely German Idealism, Neo-Kantianism and Phenomenology. We often read Kant through the lens of the appropriations of Kant effected by each of these traditions, and then blame him either for not yet reaching their stage or for sharing their failings. The German Idealists placed great stress on systematicity and unity, and blamed Kant for stopping short of this unity. When French philosophers in the 1960s came to criticize this quest for unity in Hegel, they immediately assumed Kant was equally guilty for having necessarily led to Hegel. They usually did not come to question whether this from Kant to Hegel narrative had any legitimacy outside of the philosophical historiography of philosophy specific to – Hegel. Husserl, on the other hand, attempted to develop phenomenology as a science of transcendental consciousness. This committed him to a form of Cartesianism about consciousness, whilst rejecting psychologism, which, for Husserl, meant embracing Platonism. When critics of phenomenology like Foucault and Derrida ultimately came to attack the

peculiar nexus of the transcendental and the psychological, the ideal and the concrete, in Husserlian phenomenology, they equally assumed that Husserl was somehow entitled to claim that he derived this “transcendental motif” from Kant. Finally, like Husserl, the Neo-Kantians of the *Marburger Schule* assumed that the only defense against psychologism was Platonism, and therefore chose to defend Kant from the psychologistic recuparation by stressing those elements of his philosophy that are platonistic and rationalistic. As a result, they systematically underestimated the empiricist side of Kant’s thought, as well as his fundamental criticism of the incarnation of Platonism he recognized in Leibniz and his followers. By constructing a false dichotomy between Aristotle and Plato, between rationalism and empiricism and between psychologism and platonism, they lost sight of Kant’s attempt to evade this dichotomy. Moreover, their platonism led them to overemphasize Kant’s fascination with mathematics and mathematical physics, thereby constructing the stubborn myth that Kant was mostly an apologist of Newtonian mechanics and a philosopher of science in a postivist vein.

This explains another way in which this dissertation is perhaps too long, for it is not just a commentary on Kant – it is also a commentary on the practice of commenting on Kant. I have often tried to unveil the philosophical rationale behind what I identify as misinterpretations of Kant, not to pretend that I am free of these philosophical biases, but rather to show that it is *always* a matter of philosophical argument, and never just one of historical adequacy. My question is of course the following: “how should we read Kant”; but what the precise modality of that “should” is, is precisely the issue at stake. I am trying to reveal the philosophical stakes at work in the history of philosophy, and show that they not only can, but should, figure in our arguments. I prefer to read Kant in such a way that I can still learn from him, still be surprised by him, and be confronted with all the peculiar contingencies of our philosophical history. I am not interested in reading him in order to find there what I have been taught to expect, and in such a manner that he exhibits his correct place in the indomitable progress of philosophical thought that has led us to its provisional apex in the currently most cited articles. Nor am I interested in holding fast to a lost era in which the old interpretation of Kant was held as truth, which is a nostalgia to a time that never was.

The dissertation is both too long and too short in yet another way, because of its holistic manner of argument. In the course of the seven chapters, I offer a variety of interpretations of Kantian doctrines which I admit to be controversial. My hope has been to support these interpretations not only by showing how they figure into a revised picture of the “historical” background, but also to show how they harmonize with the text and with each other. I have relied heavily on detailed analyses of whole passages, rather than on paraphrases. This is because I have often found the precise wording surprising and puzzling. There are many passages in Kant which even non-experts know well, to the extent we no longer seem to need to read them to know what

they say. I have drawn attention to these passages to show that, on closer inspection, they do not quite say what we have been taught to say, and that their phrasing and terminology often reveal resonances that we would not otherwise have expected. Paraphrasis glosses over these specificities, thereby precipitously equating terms and overlooking connections and disconnections.

Moreover, I have tried to show that the various peculiar passages resonate with each other, rather than form a series of exceptions. A major reason for the misinterpretation of Kantian arguments is that we read them not only against a supposedly stable historical background, but also against a supposedly stable background that is the *rest* of Kant's doctrines. I have tried to show that, here as elsewhere, holistic revisions are a real option. Some valuable interpretations might be rejected all too readily because they conflict with what we suppose to be Kant's official doctrine. On closer inspection, however, the fact that this background *is* Kant's official doctrine is itself questionable, not just on textual grounds, but also because it is mostly proven by the fact that it grounds the former misinterpretation.

This also reveals what I take to be a manner in which this dissertation is too short, for it would be complete only if it offered a comprehensive interpretation of the whole of Kant's theory and its relevant background. The reason why I have not done so is of course obvious, for this task would be far beyond the scope of one work, even beyond the scope of one oeuvre. The adequate philosophical evaluation of the revised picture is yet another task that goes beyond this work in several respects. I can only hope that my suggestions here and there make intelligible how I would now go about dealing with future challenges. In the conclusion, I identify some of these future tasks.

And yet, in the spirit of honesty I would like to admit that, in its current and final form, the dissertation still only fulfills half of the task I had set myself when I first conceived it in 2008. Then, I had hoped not just to show *that* the system of faculties is properly understood in analogy with an organic system, and *that* it therefore has importantly dynamical properties – I had also hoped to show *how* this translates into the argument of the transcendental deduction and *how* it explains the nature of the categories. Throughout the present work, you will find hints of this original project, and I have also delivered papers which are properly framed in it. Nevertheless, I had to omit this part of the research entirely in order to have any hope of ever finishing, for not only did the latter half of the assignment prove too daunting, the first half turned out massively more demanding than anticipated as well.



## List of Abbreviations

All references to Kant, with exception of the Critique of Pure Reason, are to the pages and volume numbers in *Königlichen Preussischen Akademie der Wissenschaften* (ed.). *Akademie Ausgabe of the Works of Immanuel Kant* (29 vols.). Berlin: G. Reimer, abbreviated AA, followed by volume number in roman numerals.

The Critique of Pure Reason is cited by *CPR*, followed by the page on the first (A), and/or second (B) edition.

Unless indicated otherwise, all English translations of passages from Kant are from Guyer, P. & Wood, A. W. (eds.) *Cambridge Edition of the Works of Immanuel Kant*. Cambridge, Cambridge University Press.

References to Descartes are from Adam, C. & Tannery, A. (eds.). *Oeuvres de Descartes* (11 vols.). Paris: Vrin, abbreviated by AT, followed by volume number in roman numerals. The sources of the English translations are noted in the reference in the text.

References to Leibniz are from the following editions, unless the manuscript was not included in them:

Gerhardt, K. I. (ed.) *Die philosophischen Schriften von Gottfried Wilhelm Leibniz*. Georg Olms Verlag, abbreviated by G, followed by volume number in roman numerals. The sources of the English translations are noted in the reference in the text.

Gerhardt, K. I. (ed.) *G. W. Leibniz: Mathematische Schriften*. Georg Olms Verlag, abbreviated by GM, followed by volume number in roman numerals. The sources of the English translations are noted in the reference in the text.

References to Aristotle are to pages and columns in the Becker edition. All translations are from Barnes, J. (ed.) (1984). *The Complete Works of Aristotle* (2 vols.). Princeton: Princeton University Press, unless otherwise indicated.

References to Malebranche are to Robinet, A. (ed.) *Oeuvres complètes de Malebranche*. Paris: Vrin, abbreviated by OC, followed by volume number in roman numerals. The sources of the English translations are noted in the reference in the text.

References to Plato are to the pages and columns of the Stephanus edition. All translations are from Cairns, H. & Hamilton, E. (eds.) *The Collected Dialogues of Plato*. Princeton: Princeton University Press.



# Table of Contents

Introduction .....	1
<i>Part 1: Dynamics, Epigenesis and History .....</i>	<i>11</i>
<b>Chapter 1      Extending Science: The Dynamic in the Pre-Critical Period.....</b>	<b>13</b>
1.1    “Nothing in Kant makes Sense except in the Light of Newton”? .....	14
1.1.1    The Emergence of Positive Science .....	17
1.1.2    Kant as Scientist.....	23
1.1.3    The Looseness of Fit between Transcendental Philosophy and Natural Science .....	26
1.1.4    Kant and the Inexact Sciences .....	34
1.1.5    The Exception as Rule .....	37
1.2    Finding the Dynamical in Dynamics .....	39
1.2.1    Sufficiency and Continuity in Leibnizian Dynamics .....	39
1.2.2    The True Estimation of Leibnizian Dynamics .....	45
1.3    From Cosmogony to Chaosmogony .....	52
1.3.1    The Problem of Cosmogony in Early Modern Science .....	52
1.3.2    Newtonian Cosmology? Nomology and Analogy .....	55
1.3.3    Time and Teleology .....	59
1.4    Conclusion .....	65
<b>Chapter 2      Epigenesis and the Dynamic Nature of Life: Kant’s Theory of                     Generation .....</b>	<b>67</b>
2.1    Kant and Early Modern Theories of Generation .....	68
2.2    Structure and Genesis in Early Modern Theories of Generation.....	71
2.2.1    Epigenesis and the Circular Paradigm.....	73
2.2.2    The Disunity of Preformationism .....	78
2.2.3    Metamorphogenesis.....	85
2.2.4    Epigenesis and the observability of structure and genesis .....	90
2.3    Epigenesis in the pre-critical period.....	94
2.3.1    The Only Possible Account of Animal Generation.....	94
2.3.2    Races and Dispositions.....	100

2.4	Epigenesis in the Critical Period.....	106
2.4.1	Germs and Dispositions as Internal Limitations.....	106
2.4.2	Generic Preformationism and Transcendental Epigenesis.....	108
2.5	Conclusion .....	111
<b>Chapter 3</b>	<b>Epigenesis and History .....</b>	<b>113</b>
3.1	Epigenesis Beyond Generation .....	114
3.2	Epigenesis and Transcendental History .....	117
3.2.1	Rousseau’s Pristine Man.....	117
3.2.2	Kant’s Pedagogy.....	121
3.2.3	The Myth of the Origin .....	125
3.3	Epigenesis and the Transcendental History of Reason.....	132
3.3.1	Kant’s Epistemological Preformationism? .....	132
3.3.2	The Epigenesis of Reason .....	138
3.4	Transcendental Idealism and Historicity.....	147
3.5	Conclusion .....	149
 <i>Part II: Organics, Systematicity and Purposiveness .....</i>		 <i>151</i>
<b>Chapter 4</b>	<b>The Structure of the Organic: Natural Purposiveness and Circularity .....</b>	<b>153</b>
4.1	From Natural Teleology to Natural Theology .....	155
4.1.1	The Naturality of Teleology.....	156
4.1.2	The Mechanization of the World-Picture as the Artifactualization of the World.....	166
4.2	Natural Purpose and Circular Causation.....	171
4.2.1	Natural Purposiveness Delineated.....	171
4.2.2	Natural Purposiveness Exemplified.....	176
4.2.3	Natural Purposiveness Determined.....	185
4.3	Assuaging the Antinomy of Teleological Judgment .....	199
4.3.1	Presentation of the Antinomy .....	199
4.3.2	Candidates for Heuristics .....	204
4.3.3	The Resolution of the Antinomy .....	210
4.4	Conclusion .....	223
<b>Chapter 5</b>	<b>Systematicity and the Disunity of Reason.....</b>	<b>227</b>
5.1	The Problem of Transcendental Systematicity .....	228
5.2	The Spirit of Systems .....	231
5.2.1	The 17 <sup>th</sup> Century Spirit of Systems .....	231
5.2.2	The Spirit of Systems under Attack.....	237
5.2.3	The Spirit of Systems on the Defense.....	245
5.3	The Disunity of Reason in the Critique of Pure Reason .....	251
5.3.1	Logical Systematicity .....	251
5.3.2	Transcendental Systematicity.....	258

5.4	The Disunity of Reason in the Critique of the Power of Judgment .....	262
5.4.1	Reflective Judgment and its Principle .....	262
5.4.2	The Antinomy of the Teleological Power of Judgment and the Antinomy of Systematicity .....	270
5.5	Conclusion .....	272
<b>Chapter 6</b>	<b>The Disunity of the System of Faculties.....</b>	<b>275</b>
6.1	Against Kantian Dualism: the Hegelian and Heideggerian Legacies .....	276
6.2	Faculty Psychology in German Rationalism .....	284
6.2.1	Wolff and the Substantial Unity of the Representative Power .....	284
6.2.2	Crusius and the Interanimation of Mental Powers .....	287
6.2.3	Tetens on the Development of the Soul.....	292
6.3	The Disunity of the Gemüt .....	300
6.3.1	The Distinctness of Sensibility and Understanding .....	300
6.3.2	The Paralogisms and the Unity of the Soul .....	310
6.3.3	The Disunity of the Faculties and the Regulative Ideas.....	323
6.4	The System of Faculties as an Organic Unity.....	330
6.5	Conclusion .....	333
 <i>Part III: Systematicity and its Historicity.....</i>		<b>335</b>
<b>Chapter 7</b>	<b>The Epigenesis of the Faculties.....</b>	<b>337</b>
7.1	Aesthetics and Cognition in the Eighteenth Century .....	337
7.2	The Functional Unity of the Faculties .....	347
7.2.1	Reflective Judgment.....	347
7.2.2	The Harmony of the Faculties and the Harmony of the Beautiful.....	352
7.2.3	The Pleasure in Harmony .....	364
7.3	The Sublime Vocation of Man.....	370
7.4	The Epigenesis of the Faculties.....	375
7.5	Conclusion .....	378
 <b>Bibliography.....</b>		<b>381</b>



# Introduction

The topic of this dissertation is the possibility of the historicization of the a priori. In other words, it is a quest for what has been, for quite some time now, a philosophical holy grail. I liken it to the holy grail for two reasons: first of all, it is generally assumed that a solid account of historicizing the a priori is capable to perform philosophical miracles, such as resolving some of the deepest tensions and struggles underlying the field as a whole; secondly, those very miraculous properties raise justifiable doubts regarding the possible existence of this philosophical panacea.

I will not pretend to have found it, but I do believe to have charted a region where our search may begin afresh, with more than vague rumours and inarticulate suspicions to go on. Or rather, I have recharted a region that has been charted many times before, and where we have already sought diligently. But I have reasons to believe we have spent that time looking in the wrong way.

The region I have recharted is the critical philosophy of Immanuel Kant. Kant is often accredited with one of the best possible accounts of the a priori, namely in terms of conditions of possibility. This account is powerful and alluring because it does not obviously locate the a priori in some transcendent realm of eternal truths, in the mind of an omniscient and omnipotent creator or even in the supposed core structure of the world – groundings that are dissatisfying to the austere tastes of many of us. Rather, this a priori seems to flow directly from the acts of thinking and knowing, and is therefore embedded in a finite, dynamic practice of understanding.

Yet, all of this raises the question as to *why* these and no other conditions constrain our thought. Kant has usually been regarded as having provided only a partial answer to this question, for he can at best be read to succeed in showing that these constraints are conditions of the possibility of knowledge. But it is far less obvious whether a good argument can be extracted from his writings to support the view that they are *the* conditions of possibility, in that there can be no other. Quips concerning the supposed arbitrariness of the table of the categories abound in the philosophical literature after the publication of the *Critique of Pure Reason*, as do sneers that Kant is simply elevating

the principles of his own, scientifically and philosophically blinkered, epoch to eternal truths.

These charges of arbitrariness and partiality have led many to suspect that there may be other systems of categories, and that there may even *have been* other systems of categories. In fact, the idea that Kant's philosophy is somehow irremediably marred by its commitment to the rigidity of the a priori is shared by most of the major movements of 19<sup>th</sup> century German philosophy, which are almost universally either post-Kantian or neo-Kantian.

## Dynamizing the A priori

The first receptions of Kant's philosophy took issue with the dualisms of his thought, which are supposed to be a priori and fixed. Most Post-Kantian German philosophers were somehow involved in the project of overcoming Kantian dualism and the bruteness of the a priori. To this end, the philosophies of the early Fichte and Schelling meant to replace the ready-made, absolute positing of the faculties and the categories with a genetic account in terms of an absolute unity coming to differentiate itself in specific ways through an internal process of self-limitation. The specificity of the categories and the faculties can then be explained by showing how they can come to be through a process the form of which is necessitated by the type of original unity involved.

This genetic method has entered philosophical common sense mostly through its elaboration by Hegel. Hegel pointed towards dialectics and the negative as the major moments of such a process. In addition, Hegel's philosophy is widely believed to be original because it no longer wishes to justify the specific faculties and categories as necessary through a reconstruction of their genesis, but argues instead that the faculties and categories are necessary only relative to a certain stage in the self-development of spirit and its knowledge of itself. In Hegel's view, the Kantian a priori's are indeed necessary to a certain stage in the development, but the dualisms and bruteness they exhibit is also the reason why this stage must end up in a contradiction. This contradiction is eventually solved through a dialectical sublation where the previous stages are negated, but in such a way that what is true and necessary about them is preserved in the new stage. The Hegelian philosophy aims at developing the picture of knowledge that would result from the final stage in this dialectical process, namely absolute knowledge.

Contrary to what the many horribly inadequate historical overviews of philosophy of the 20<sup>th</sup> century would have their readers believe, German Idealism did not reign

supreme in German philosophical circles in the 19<sup>th</sup> century. From relatively early on in the century, there was a current of neo-Kantianism that fought against the purported attempts of Post-Kantians to reinstate metaphysical speculation as a legitimate, and perhaps *the proper* business of philosophy. Instead, they wanted to keep reflecting on the necessary limits of human knowledge insofar as they are expressed in the Kantian a priori's. But like many readers of Kant, they believed that the reasons why these and precisely these a priori's constituted the limits of human knowledge should be explicated beyond Kant's own summary treatment of the issue. Over the first half of the 19<sup>th</sup> century, a position developed that we now know as psychological or physiological neo-Kantianism. Its proponents held that Kantian a priori is not necessary for being logically or metaphysically necessary, but for being the general laws governing human perceptual and psychological experience. The a priori is innate and fixed because it is ultimately a biological and/or psychological fact about the human species. There may also be mixed positions, which borrow from the early German Idealists the idea that the a priori is linked to genetic processes, but that maintain that these geneses are not of metaphysical entities, but rather of concrete psychological individuals and their store of percepts and concepts.

By the late 1860s, as physiological neo-Kantianism had risen to dominance in German academia and started professing that epistemology or theory of knowledge should be the foundation of all philosophy and should be concerned with the investigation of how we think, i.e. how our psychological and physiological apparatus work, a new movement of neo-Kantianism arose. This movement, whilst sharing their predecessors' reticence towards metaphysics and their preference for theory of knowledge, railed against the psychologism of earlier neo-Kantianisms. A crucial figure in this movement is Hermann Cohen, who sought to distinguish Kant's own reflections on the a priori from the subsequent psychophysiological projects of Herbart and his epigones. He protested that this interpretation of epistemology could only give us information about how men tend to think, about the psychological and physiological mechanisms that give rise to our beliefs and representations. But this has no bearing on the justification of that knowledge. To use Kant's image, the empiricist accounts of how we tend to arrive at knowledge only settles the matter of the facts, not that of rights.

Cohen suggested to reform theory of knowledge, which is concerned with the principles of our empirical processes of reasoning, as Critique of knowledge, which aims at identifying the grounding assumptions of our best scientific theories. He suggested that Kant was engaged in the latter enterprise because he intended to identify the underlying principles, the conditions of possibility, of Newtonian Science. Theory of Knowledge, for Cohen, had both a new target and a new object: it was no longer concerned with investigating the general regularities that govern thought, but with the fundamental principles that underlie thought; and it was no longer interested in knowledge as a psychological phenomenon, but in knowledge as scientific theory and

investigation. In the early Marburger Schule, real knowledge, justified knowledge, came to be identical with science, and science with our best theories.

As fascinating as this attempt is, it was immediately followed by a criticism. Many were tempted to accept Cohen's interpretation of Kant's critical philosophy as correctly identifying the a priori of Newtonian Mechanics and Euclidean Geometry, but then continued to point out that both of these scientific theories were on their way out. Over the course of the 19<sup>th</sup> century, a tendency towards algebraization and formalization had changed the practices and theoretical constraints of mathematics irrevocably. A crucial consequence of this is that the Euclidean parallel postulate was no longer considered to be necessary, but was rather viewed as a characteristic of only one in an infinite number of possible alternative geometries. Secondly, the fundamental concepts and methods of mathematics were no longer derived from the method of pure construction prescribed by Euclid's Elements. By the early 20<sup>th</sup> century, moreover, Newtonian Mechanics came to be challenged by Special and General Relativity, both of whom denied that Euclidean Space is even proper to characterize the spatio-temporal grid of our empirical theories (the so-called Einstein-Minkowski space). Some philosophers took this to mean that there could be no a priori, and that any and all so-called a priori principles are defeasible on empirical grounds.

This was not to deter the Marburger Schule, however, since by then it had admitted that the changes in our science also bring along changes in our fundamental principles, even though they believed that some a priori principles remained in place. This theory was most famously articulated by Ernst Cassirer, who admitted that the principles of Newtonian Mechanics can be rejected, but maintained that the very fact of the revisal of our knowledge entails that there are still more general principles governing our thought. The a priori principles of science were now no longer thought to be those of our best scientific theories and practices, but those of the continuous and dynamic process that is science in general. And yet, there are second-order a priori principles that are necessitated by a specific stage of any science – Kant's categories, for instance, correctly grasp the supra-empirical framework of Newtonian mechanics, even though they need to be abandoned together with the framework it makes possible.

Both the German Idealist and the Marburger Schule conceptions of a dynamical a priori have influenced many approaches in the 20<sup>th</sup> century, and the reader will forgive me if I abstain from tracing its influence through such varied movements as positivist, logical positivist and post-positivist philosophy of science, as well as French historical epistemology and the structuralist and post-structuralist philosophies that engage with it. I will rather go straight to what I take to be the defects of both accounts, defects that they risk passing on to their progeny.



## Locating the Dynamics of the A Priori

Although the post-kantians and the neo-kantians essentially agree that a dynamic element is required to somehow validate Kant's list, or rather table, of a priori principles, they disagree on how to conceive of this dynamics. Because of its positivist strand, the Marburger Schule neo-kantians seem to believe that the locus of dynamics lies in science, narrowly conceived as a given epoch's best theoretical and institutional framework of inquiry into nature. Various assumptions lie behind this. First of all, there is the belief that there is a distinct, isolable and recognizable subset, a natural kind if you will, of human practices that constitutes science. Secondly, there is the conviction that this thing called science has a higher epistemic status than non-scientific practices, and is thus normatively elevated above the latter, to such an extent that it provides the sole standard of the epistemic worth investigating. As a result, theory of knowledge is concerned only with the a priori principles governing theories and scientific frameworks or traditions, not with cultural practices or epistemic practices the scientific status of which is all too disputed. Moreover, the dynamics that epistemology must admit is the dynamics of theory-change, of tradition-change, of paradigm-change, etc. in science, not the changes of anything more socio-culturally general.

This science-centrism is typical of much late-19<sup>th</sup> and early 20<sup>th</sup> century philosophy, and it is a clear shortcoming of it. First of all, its implicit or explicit devaluation of non-scientific epistemic practices leads to the bizarre result that non-scientists do not actually think and have no a priori principles, unless one assumes that they simply engage in a lower grade of science, or folk science, or adhere to outdated scientific frameworks. The idea that most people will never truly be engaged in "knowing" seems to me incredibly arrogant and even preposterous.

There is a reason for this bizarre result, and it is the legacy of positivism in 20<sup>th</sup> century philosophy. However critical philosophers may position themselves towards this tradition, they often continue to work within its general framework, which is that philosophy should primarily be engaged with science as the crowning human ratiocinative accomplishment. This focus on science persists even when philosophers realize, or should realize, that they are engaged with something much broader. As a result, our thought about conceptual schemes is usually in terms of scientific theories.

This science-centrism sets the agenda for the question of the dynamization of the a priori as well. For many, it is the question of whether science can have the epistemic pride of place credited to it by positivism, and whether it can also be said to *progress*. After all, according to an older conception, science derives its pride of place from its high conformity with the a priori principles governing knowledge. But if these principles turn out to be *relative to* science, they can no longer independently justify it. The question then becomes that of how science can still lay claim to its epistemic

authority. Additionally, many see the virtues of science in the fact that it leads to progress in our knowledge. Progress, however, seems to require a continuity of the standard with which our body of knowledge is evaluated. This becomes problematic once the standard becomes relative to a given stage of science. The challenge then becomes to describe the dynamic of science in such a way that the latter framework preserves the legitimacy of the former and insures that the latter constitutes a recognizable advance on the former.

This concern over the privilege of science is not mine, for I do not believe we can justifiably accord science the pride of place it holds in our society, since there is no justifiable way to characterize science such that it is revealed to be a distinct and privileged epistemic practice. The many things we call science, and which we have called science, do not constitute a more than remotely recognizable natural or even social kind, and there is no recognizable general framework governing them all. This is not just because science is diverse and plural rather than one and unified, it is also because science cannot be clearly delineated from what some would like to call unscientific, pre-scientific or even pseudo-scientific practices. Thus, not only is scientific understanding not the sole or the highest form of understanding, it is not even a recognizably distinct form of understanding.

The result is that we need a broader conception of understanding, knowing and judging than the one for which the positivist legacy allows. This also means that we need to look for the conditions of possibility of this broader sense of knowledge, and not just for those of the unjustifiably narrow sense of knowledge. At first sight, the German Idealist, and specifically the Hegelian route seems more promising for such an enterprise. After all, the stages of history examined by Hegel are not just stages in the progress of science: they are stages in the progress of human epistemic, moral, aesthetic, social and cultural life.

Nevertheless, I have serious qualms with the Hegelian model as well, qualms which I have inherited from the critical reception of Hegelianism in French philosophy in the latter half of the previous century. The two major problems are the tendency to simplify the motor of the dynamics of the a priori, and the tendency to regard this dynamics as predestined, tendencies that I believe to be importantly related.

The Hegelian conception of dynamics is the process known as the dialectic, and its motor is the negation or the contradiction. Each stage of the development of Spirit has a principle which constitutes its essence and its legitimacy. This principle, however, inevitably includes an essential contradiction. Being essential, this contradiction cannot simply be appeased, and must eventually come to the surface and lead to change. This change is a dialectical sublation of the contradictory pair in a new principle, which becomes the principle of a new epoch. The underlying belief thus seems to be that, for every epoch, there is one essence, and one constitutive contradiction, and that it is this contradiction that leads to the new stage.

This unicity of the principle and of the contradiction has deterministic consequences, since the eventual change is perfectly determined by the fact that there is only one contradiction which can lead to change, and there is only one solution to that contradiction. Add to this the further assumption, which one can arguably ascribe to Hegel, that there is one and only one starting stage of humanity, and one and only one end-stage, then we have a preset course of events, which can change in details, but not in respect of the constitutive contradictions.

The attractive features of this model are clear: it can regard the principles of each stage as necessary because they are necessitated by the necessary historical development of rationality in a broad sense. Thus, although they have their limitations and will be overcome, they are still legitimate within their own time and place. Also, although the actual change may be prompted by historical contingencies, these historical contingencies only prompt this change because they have the structure that is predetermined to prompt change. Finally, the dialectical process guarantees that every subsequent change preserves the validity of the previous stage, albeit in a now restricted manner.

My dissatisfaction with this picture, however, stems from my feeling that it is of a mock-historical nature. Although reason necessarily goes through temporal changes, these changes were all necessitated in advance, and are all directed towards a clear identifiable end-stage. If by “historical” one means “temporal”, this is perhaps satisfying, but if one means by it “open to historical contingency”, then it is clearly dissatisfying. After all, historical contingencies can only occasion predetermined changes, which means that the contingent features of the contingency play no role in the change – only those features that match the necessity do.

This charge raises the stakes for a proper account of the historicization of the a priori, for it not only demands that the a priori maintains its legitimacy even in the face of its temporally evanescent nature, but also that it maintains it in the face of its openness to historical contingency. My goal in this dissertation has been to offer a reading of Kant that provides the basis for such an account.

## **The Organic and the Dynamic**

In this dissertation, I suggest that we can arrive at a historicized and dynamized a priori by interpreting the Kantian system of faculties as an organic system. Kant himself uses two important metaphors from the life sciences to characterize this system, namely the idea of an “epigenesis of reason”, and his likening the system of faculties, and other systems, with the living system as organized. These two metaphors relate strangely: the

very idea of an epigenesis of reason suggest a developmental and historical dimension to the a priori, whereas the organic analogy invokes the kind of closure and intimate interrelation of parts that is unlikely to arise from historical contingencies. I therefore seek to dispel this apparent contradiction between the dynamic and the systematic implications of the organic metaphor by merging them into a coherent picture of the dynamic nature of systematicity.

In part I, I develop my reading of Kantian epigenesis, suggesting that it is both central to Kant's thought and more dynamic than usually presumed. Chapter 1 is meant to show that Kant's concern with dynamic systems and with epigenesis is not an afterthought that stands in tension with his mostly mechanical interests. Instead, I argue that the issue lay at the heart of Kant's thought on what we would now characterize as physical topics as well. What Kant was most concerned with in these topics, I submit, is exactly how dynamical dynamics really is. His conclusion seems to be that Leibniz and Newton fail to provide a good basis for a dynamic understanding of nature.

This tension between mechanics and the dynamic is also at the heart of Kant's reflections on early modern embryology, which form the topic of chapter 2. There, I argue against the suggestion of recent commentators that Kant's version of epigenesis was ultimately far less dynamic than its name suggests. I argue instead that Kant intended epigenesis to be a genuine alternative to both metamorphosis and preformation. The resultant theory insists that the dynamics of an organism is properly understood internally to that organism, but that it nonetheless has an internal openness to contingency.

In the third chapter, I tackle Kant's usage of embryological language in his philosophy of history and epistemology. The reading of Kantian epigenesis as ultimately preformationist, and therefore non-dynamical, which I attacked in chapter 2, has led some commentators to deny real dynamics in Kant's philosophy of history and epistemology as well. I argue against such readings by showing that Kant's usage of embryological metaphors in these other parts of his philosophy are better understood in the light of the more dynamical picture of epigenesis I have sketched in chapter 2. The result is that, for Kant, the history of human ratiocinative capacities is much like the developmental history of an organism in that it is governed by internal dynamics that are nevertheless open to external contingencies.

The result of part I, however, requires that we gain a better understanding in the kind of system that Kant believes to be subject to internal dynamics and yet open to contingency. In part II, I develop my interpretation of Kant's conception of such a system.

In Chapter 4, I offer my interpretation of Kant's concept of a natural purpose. I argue that a natural purpose is a normative unity of a diversity rather than a substantial unity, and that this unity is autonomous and plastic. I also argue that natural purposes are

incomprehensible to us because our cognitive system consists of two distinct functions that do not automatically harmonize, namely sensibility and understanding.

In Chapter 5, I then argue that this potential for disharmony is present in the very nature of reason, in the form of an antinomy between two drives of reason. My main goal there is to show that the drive for unity is not the sole force behind Kant's conception of the dynamics of our conceptual system: there is a drive towards multiplicity as well. The goal is then to harmonize these two drives, not to sacrifice one of them to the other. Reason itself thus appears as a normative unity of diverse and even possible conflicting drives.

In Chapter 6, I argue that not just reason, but the system of faculties as a whole constitutes a normative unity of the diverse, in the sense in which an organism is supposed to be as well. The different faculties are like so many organs, and the proper functioning of our cognition depends on their proper alignment and cooperation just like the proper functioning of the organism requires its physiological harmony. Thus, the system of faculties is open to contingency because it is diverse – after all, whether the harmony obtains or not can only appear in every new cognitive situation.

In chapter 7, I argue that it is even possible to read Kant as allowing for the plasticity of the system of faculties. Just like an organism can reconstitute itself in the face of historical contingency, so the cognitive system can adapt to new situations upon the realization that its unity systematically fails to obtain. I argue that the motor behind this plasticity is reflective judgment, and that the feelings of the beautiful and the sublime play the role in the system of faculties that pain and pleasure play in organisms.

I conclude, then, by sketching the picture of the system of faculties as a dynamic and plastic system open to contingency, and listing some consequences. One consequence is that the categories may be determined not just by the laws of logic, but by the relation of the faculties. If this is true, a change in the relation of the faculties or in the nature of the faculties itself (which often comes down to the same thing), would imply a change in the categories. We can thus have a theory of why the categories have their specific form and how they could come to acquire a different form. It is also important that we can understand the normative force of the categories in spite of their dynamic and somewhat contingent nature, for this normative force derives from the fact that the unity of the faculties is a normative unity.

Another is that it is reflective judgment that forms the motor of dynamics, and that we should therefore acquire a better understanding of Kant's theory of judgement, and of the notion of judgment in general. By linking the a priori to a general notion of judgement, which can be cognitive, moral, aesthetic or even legal, rather than to science or scientific theories, we can transcend science-centrism. Our task is then to understand judgment, in the many different forms it takes both synchronically and diachronically, rather than knowledge, in the form it has taken since the dawn of Modernity in Europe.

I also list some vistas of research opened up by arguments that have popped up in the course of the main argument of the dissertation, and finally comment upon the proper modality of its claim. I will not do so here, partly in fear of eventually repeating myself later on, and partly because I feel that the introduction can only offer a general sketch of the problem and of the course of argument. No substantial claim I would make here would be properly intelligible until after the work itself has been read. For this reason, I can only request your patience.

**Part 1:**

**Dynamics, Epigenesis and History**





## Chapter 1

# Extending Science: The Dynamic in the Pre-Critical Period

*Die Tradition der Unterdrückten belehrt uns darüber, daß der "Ausnahmestand", in dem wir leben, die Regel ist. Wir müssen zu einem Begriff der Geschichte kommen, der dem entspricht. [...] Das Staunen darüber, daß die Dinge, die wir erleben, im zwanzigsten Jahrhundert 'noch' möglich sind, ist kein philosophisches. Es steht nicht am Anfang einer Erkenntnis, es sei denn der, daß der Vorstellung von Geschichte, aus der es stammt, nicht zu halten ist.*

- Walter Benjamin

In this chapter, I will argue that Kant's engagement with and project for physics centres on the problem of teleology. This goes squarely against the wide-spread view of Kant as torn between his enthusiasm for modern (positive) science, with its prohibition on teleology, and his allegiance to the teleological picture suggested by theology and the (at the time) less well-grounded sciences. My argument involves a negative and a positive phase. The negative phase consists in a sustained criticism of the traditional picture of Kant's attitude to physics, and the role its exemplar, (Newtonian) mechanics, fulfills within the transcendental philosophy. In order to do this, I will identify, in section 1.1, the four main pillars on which this traditional picture rests, subject each of these pillars to criticism (in subsections 1.1.1, 1.1.2, 1.1.3 and 1.1.4) and draw from this (in subsection 1.1.5) the conclusion that an alternative picture of Kant's position towards physics is suggested by his specific interest in the physical subdomains of dynamics and cosmogony. The next two sections of the chapter are devoted to developing such an alternative picture from an interpretation of Kant's earliest engagements with both dynamics and cosmogony. In section 1.2, I argue that Kant's

interest in dynamics was prompted by the realization that it is in tension with mechanics and the mathematical approach it countenances because of the irreducible involvement of time and temporal development in its framing of causality. In order to make this plausible, I first sketch some important features of Leibnizian dynamics that could have suggested this puzzle to Kant (in subsection 1.2.1) and then offer an interpretation of his treatment of dynamics in the 1740s and 1750s, which yields the conclusion that Kant was struggling to uphold the causal link in dynamic interactions, without thereby eliminating genuine change and evolution from our descriptions and explanations of natural systems (in subsection 1.2.2). In section 1.3, I propose that a similar concern can be discerned in Kant's theory of cosmogony from the 1750s. I first offer (in subsection 1.3.1) some background to the problematic status of cosmogony in early modernity, in order to identify the expectations and worries that lay behind Kant's reasoning on the subject. Next, I submit (in subsection 1.3.2) that Kant was inspired, in this early work, not only by the content and ambition of Georges Buffon's attempt a theory of the formation of the solar system, but also by the great French Natural Historian's antimathematicism and analogy-driven reasoning. Finally, I argue (in subsection 1.3.3) that his work on cosmogony led Kant to a hesitant position towards teleology, and a first attempt to distinguish different kinds of teleological reasoning and explanation.

## 1.1 “Nothing in Kant makes Sense except in the Light of Newton”?

It is hardly a *bold* conjecture that modern Kant scholarship started with Hermann Cohen's brilliant reinterpretation of the transcendental philosophy, expressed programmatically in his 1883 Lecture *Von Kants Einfluß auf die deutsche Kultur*:

The kantian philosophy is indeed, in its theoretical part, first and foremost nothing else than the legitimation of these physical exemplars, the proof of their cognitive value in connection with *natural science that rests on mathematics*. Such a proof is the deed of *philosophical genius*. In this only has philosophical genius proven itself everywhere, in Plato and Descartes, in Leibniz and Kant, that it has asked the question: *What is Science?* Kant's philosophical advantage over his predecessors can be determined as follows: that whilst Descartes and Leibniz were simultaneously contributing to the establishment of science, *his* force was

concentrated on that one philosophical question. For since Newton science had become an articulated reality. (Cohen 1883: 7-8; original stress; my translation)

In this interpretation, which Cohen had advanced 13 years before in his monumental *Kants Theorie der Erfahrung*, Kant surfaces as the champion of a scientifically informed but theoretically autonomous epistemology, rather than as a predecessor to the great idealist systems of the early nineteenth century and the psychologicistic and physiologicistic enterprises of the middle of the nineteenth century. He does so, however, only insofar as he is engaged in legitimating natural science, and a particular system of natural science at that: “Kant’s task is hence first and foremost the scrutiny of and characterization of the cognitive value and foundation of certainty of Newton’s natural science, that he grasped through the thread of Experience. [...] The transcendental method originated in the reflection on the *Philosophiae naturalis Principia mathematica*.” (Cohen 1885: 66-67; my translation). As a result of the efforts of Cohen and his followers, the so-called *Marburger Schule* of neokantianism, the names Newton and Kant, that of the British epitome of stern scientific enquiry and that of the German herald of stern philosophical critique, would become intimately intertwined. In spite of the many criticisms of and deep dissatisfactions with the *Marburger* rendition of the critical philosophy, hardly anyone has sought to cast doubt on Kant’s alliance with and allegiance to the Newtonian paradigm that had arduously acquired supremacy during his lifetime<sup>1</sup>. In fact, Cohen’s approach was given new impetus at the end of the 20<sup>th</sup> century when Michael Friedman picked up on the *Marburger* account and modernized it in a fascinating manner. Much like Cohen, Friedman thoroughly distinguished Kant’s epistemological project from the scientific one, as is evident from the statement that he does

not mean to suggest, however, that the Kantian philosophy can be thereby seen as wholly parasitic on the exact sciences-so that, for example, one can simply read off the content of that philosophy from the scientific developments in question. On the contrary, Kant’s achievement consists rather in adapting and radically

---

<sup>1</sup> To reinforce this claim, one can for instance point out that Robert Butts was hardly abusing his licence as editor of the major volume on *Kant’s Philosophy of Physical Science* when he wrote that “[a]ll of the papers provide clear emphasis on the fact that throughout his career-from the publication of NTH in 1755 to the appearance of MAN in 1786-Kant regarded the physical synthesis worked out by Newton as the best example we have of reliable theoretical knowledge, and at the same time as the best example we have of justified method in science and metaphysics. And so there seems no disagreement over the fact that the actual physical science for which Kant attempts to provide additional philosophical credentials is Newtonian science.” (Butts 1986: 20) Moreover, Butts is hardly tacit on the source for this view: “We must look to one of the rescue boats put afloat by the Marburg neo-Kantians about one hundred years following Kant’s death for a balanced statement of the content of MAN [...] it would be difficult to find a better short statement of the main contents of MAN than this one by Cassirer” (Butts 1986: 12).

transforming independently given philosophical and metaphysical ideas-ideas stemming largely from the Leibnizian philosophical tradition he inherited-within the essentially new scientific context wrought by Newton. (Friedman 1992b: xiii)

Still, it is clear that the Kantian project is presented as essentially rather humble, as being in line with John Locke's self-characterization from the Epistle to the Reader the *Essay Concerning Human Understanding*: "in an age that produces such masters as the great Huygenius, and the incomparably Mr. Newton, with some other of that strain, it is ambition enough to be employed as an under-labourer in clearing ground a little, and removing some of the rubbish that lies in the way" (Locke 1975: 10). The "rubbish" that Kant is seeking to clear away is then the metaphysical objections that obstruct scientific advance. Indeed, to many German readers, Newton wedded the most brilliant and rigorous of scientific deductions to the sloppiest and most untenable metaphysical conceptualizations (Cf. Ahnert 2004). It would not be surprising, therefore, that the topics of Newton's *Metaphysics* – "[Newton's] discussions are largely limited to questions about the ontology of space and time, the laws of motion and the forces that cause motion, our knowledge of matter within physics, and God's relation to the physical world" (Janiak 2008: 7) – are considered to be the main topics of Kant's precritical –and in a way his critical – metaphysics.

As valuable as this position has been for recognizing Kant's acuity as a connoisseur of 18<sup>th</sup> century science, it is also, in my opinion, deeply misleading on the subject of his relation to natural science, and to Newtonian science in particular. In particular, It is based on four main pillars, each of which is doubtful in light of more recent scholarship in the history of science and Kant-exegesis. These pillars are easily identifiable in the passages quoted above:

The first pillar is the idea that Kant was taking account of the "Factum" of positive Newtonian science. It consists in the claim that in the formative years of Kant's intellectual development, Newton's work had resulted in a rift between natural science and philosophy, at least as far as their methods and research focus are concerned, and that Kant acknowledged this gap. I will argue against this claim by showing that it reads much later developments (those of analytical mechanics) back into earlier ones, and that it fails to account for Kant's interests and comments concerning Newton.

The second pillar is the central assumption that Kant differs from important predecessors like Descartes and Leibniz because in contrast to the latter, he is not also a scientist but merely a philosopher, at least in spirit, because he lacks their quantitative and experimental focus. I argue against this view by advancing that neither Descartes nor Leibniz were as unambiguously "scientific" in that sense of the word, and that, moreover, ascribing this interpretation of "scientific" to Kant is implausible.

The third pillar is the claim that Kant's transcendental analytic is meant to directly provide the basis for *Newtonian* science in particular. I will respond to this by defending Buchdahl's thesis that there is an irreducible "looseness of fit" between the principles of

the understanding and the laws of nature against Friedman's more recent attack on it, and thereby arguing that Kant's transcendental philosophy does not privilege any particular scientific theory in that manner.

The fourth pillar is the claim that Kant clearly preferred physics over other sciences. I challenge this claim by showing that Kant was clearly engaged in these other sciences, and as such had to provide some legitimation for them, albeit a different kind of legitimation than we would initially expect.

### 1.1.1 The Emergence of Positive Science

The first aspect of the *Marburger* picture of Kant's philosophical enterprise that I wish to discuss is the claim that Kant departed from the "Factum" of (Newtonian) Science (Cohen 1883: 7), i.e. from the establishment of an autonomous scientific practice that had divorced itself from metaphysics to wed itself to mathematics, a much more enabling and less obstructing partner. This view finds its origin mainly in the *Prolegomena*, where Kant states emphatically that "There is then in fact a purely natural science, and now the question is: *How is it possible*" (AA IV: 295; original stress), and moreover admits further on that

Pure mathematics and pure natural science would not have needed, *for the purpose of their own security* and certainty, a deduction of the sort that we have hitherto accomplished for them both; for the first is supported by its own evidence, whereas the second, though arising from pure sources of the understanding, is nonetheless supported from experience and thoroughgoing confirmation by it. (AA IV: 327; original stress)

These passages are certainly telling of Kant's attitude to take Science seriously in its own domain, and show clearly that he believed the sciences to be well-developed by the 1780s. Nonetheless, what is not clear here is that Kant is speaking of *Newtonian Science*. The name Newton is remarkably absent from these passages, and although Kant makes mention of "truly universal laws of nature, that exist fully *a priori*", the examples he gives are "that substance remains and persists, that everything that happens always previously is determined by a cause according to constant laws, and so on." (AA IV: 327). These are not obviously Newtonian principles, as they are intended to be more general. The first one implies merely that all change has to be understood against the background of something that remains stable, and the second that every change is caused by some other occurrence. Together, these principles at best imply that there must be some conservation principle. This assumption, however, is common to many rival-theories of Newtonian mechanics as well, and can thus hardly be interpreted as a clear reference to the *Principia's* Laws of Motion.

One could counter that, whereas the characterization given here is certainly underdetermined between different strands of early modern natural science, it is obvious from the *historical context* that by “science”, Kant means “Newtonian Science”. Now this is precisely where the story goes wrong: there seems to be a definite problem in regarding the established scientific practice of Kant’s era as Newtonian. Indeed, as Andrew Janiak states, “In the hands of figures like Laplace and Lagrange, Newton’s work led to the progressive development of Newtonian mechanics, and its practitioners embraced a conception of their discipline in which philosophical matters played little role” (Janiak 2008: 4). But the earliest efforts of Laplace and Lagrange for developing such an analytical mechanics date from the 1770s and the 1780s, well after Kant’s “conversion to Newton”.<sup>2</sup> Before this time, Newton’s ideas were still struggling to find foothold on the European mainland, and found themselves bound up with metaphysical (or apologetical anti-metaphysical) issues<sup>3</sup>. Friedman’s treatment of Kant’s precritical development is interesting precisely because he nuanced the traditional *Marburger* notion that Kant was reflecting on a well-established and largely unproblematic scientific program, and situated his ideas in the debates amongst continental Newtonians and anti-Newtonians in the mid-18th century. Whilst more historically plausible, this thesis remains problematic in the light of recent research on this debate.

The most important objection one could raise is that there was no single one tradition of “Newtonianism” at work on the continent at the time of Kant’s conversion. As it turns out, Kant’s position differs as markedly from the two figures usually cited as

---

<sup>2</sup> There are even good reasons to doubt the “independence” of Laplace’s picture of mechanics from philosophy: indeed, Marij van Strien has convincingly argued that “Laplace’s determinism was based on a re-interpretation of Leibniz’s principle of continuity, rather than on his mechanics” (Van Strien 2014: 30).

<sup>3</sup> The history of the reception of Newton on the continent is a curious one indeed. Many authors have noted the time-lapse between the publication of the *Principia* and its acceptance by scientists across the continent. Jonathan Israel’s position is characteristic of this: “until the 1730s the European reception of both Locke and Newton was so hesitant and slow as to constitute a meaningful historical problem of its own. Far from advancing triumphantly, Newtonianism was scarcely known in France before the later 1720s” (Israel 2001: 523). Of course, Israel embraces this account because it helps him show that the earlier Enlightenment could not have been Newtonian-Lockean in character, and was thus philosophically distinct from the *philosophes* High Enlightenment. J.B. Shank has recently cast doubt on this story, however, by advancing that “Partisan warfare for and against Newton’s work in the *Principia* was especially rare in the first two decades of the book’s reception, and without it the text was instead absorbed comfortably, if not always calmly, into the preexisting structures of French science” (Shank 2008: 49). This insight serves to distinguish many different ways of and motives for relating to Newton in continental science and philosophy in the early 18th century, and suggests moreover that an interest in Newton does not necessarily bring along adherence to the High Enlightenment picture of his function for the progress of science. In many ways, Kant’s relation towards Newton bears witness to the different strata of reception, and focusing exclusively on one such stratum will impede rather than deepen our understanding of it.

his main influences in these matters, Leonhard Euler and Pierre-Louis Moreau de Maupertuis, as the positions of these two differed from each other.

Many authors have referred to Euler as a major source of Kant's early thought, if not through his scientific and metaphysical views (Riehl 1908: 334; Timmerding 1918), then at least through his ideas on the relation between metaphysics and physics (Friedman 1992b; Speiser 1934: 9). This is certainly suggested by Kant's frequent approving references to them, but it cannot be accepted without some important caveats.

The first is that Euler is hardly an unambiguous representative of the turn to "positive" science, autonomous from metaphysical speculation. Of course, we find passages in his work that suggest such a picture, such as the following one:

Even if we would not be capable of proving them [i.e. the principles of mechanics] by means of the general principles of metaphysics, the marvelous conformity of all the conclusions that we can draw from them through mathematics<sup>4</sup>, with all the movements of both solid and fluid bodies on earth, and even with the movements of the celestial bodies, would suffice to put their truth beyond all doubt. (Euler 1748: 324; my translation)

But in fact, Euler seems to have adopted and discarded this attitude as it fitted him. This surfaces in his diatribe against the realist interpretation of Newtonian attraction in his *Lettres a une princesse d'Allemagne*. There, he obviously does agree with the *physical* givenness of attraction: "Since it is certain that, in considering any two bodies, the one is attracted to the other, people inquire into the cause of this mutual affinity; it is on this topic that opinions are very divided" (Euler 1842: 266; my translation); but he doesn't believe it can be accepted on *metaphysical* grounds:

The ancient philosophers have contented themselves with explaining the phenomena of the world by means of these kinds of qualities that they have named *occult*, saying for instance, that opium makes one sleep in virtue of an occult quality that makes it specific for inducing sleep; which is saying nothing at all, or rather wanting to hide one's ignorance. We should therefore also regard attraction as an occult quality, in as far as we take it to be an essential property of bodies; but since today we seek to ban all occult qualities from philosophy, attraction taken in this sense should also be banned. " (Euler 1842: 267-268; my translation)

---

<sup>4</sup> Euler speaks here of "le calcul", which could mean either mathematics in general or calculus in particular. I believe there is little reason to suppose that he is speaking specifically of calculus in this passage, and that it is more natural to take him as making a point about the use of mathematics in physics in general. I therefore chose to translate "calcul" as "mathematics".

His insisting on the reality of (absolute) space and time whilst denying the dynamical properties of matter is the result of Euler's double allegiance: to science on the one hand and to religion on the other<sup>5</sup>. As Breidert (2007: 100) puts it: "In philosophy Euler takes up the pen only in cases where he is convinced that he has to protect the Holy Bible or the sciences against philosophers' attacks or against their false doctrines."

This brings us directly to the second caveat: Kant cannot be straightforwardly regarded as embracing Euler's reflections for the rather mundane reason that he took exactly the opposite position with respect to what was acceptable in Newton's picture and what was clearly in violation of the most basic tenets of modern science. As Giorgio Tonelli (1959: 16; my translation) expressed it: "Kant inverted the Newtonian relation between space and attraction, metaphysically putting the latter before the former"; at several instances, even during his Newtonian period, Kant expressed doubts regarding the status of Newtonian absolute space and time, and argued that the latter was in fact purely relational and the effect of the real physical interaction that he often identified with attraction. Indicative of the strain on any interpretation that stresses the relation between Kant and Euler is the following statement from Peter Harman: "While Kant did not accept Euler's views uncritically, Euler's writings on the conceptual foundations of mechanics were clearly instrumental in deepening his understanding of the problems of natural philosophy, bringing about a re-structuring of the relations between the concepts of 'force', 'inertia' and 'impenetrability' in the *Metaphysical Foundations*" (Harman 1983: 240). There is, of course, a considerable difference between the claim that Kant changed his views upon reading Euler and the claim that Kant adopted Euler's views upon reading Euler. The latter claim is the stronger one, suggesting a deep agreement between Kant and Euler on many themes, whereas the former is much weaker, and simply indicates that Kant took Euler seriously. This passage from Harman reveals that there is only evidence for the weaker claim, and should hence make us wary of positions that seem to require the stronger (which are more common).

The case of that other paragon of Newtonianism in the Prussian Academy of Science, Maupertuis, is at least as puzzling as Euler's. Much like Euler, Maupertuis undertook the task of legitimating Newton's scientific enterprise to the metaphysically reticent continental literati. There is, however, an important difference between their defensive

---

<sup>5</sup> I discern two ways in which Euler's rejection of dynamical properties of matter was informed by religious concerns. On the one hand, he feared that the Wolffian tendency to regard forces as principles of change internal to the body that undergoes that change would lead to the denial of causal efficacy to not only all created substances (in fact advocated by the Wolffians) but ultimately to God as well (Euler 1842: 311-315). On the other, he saw that materialists sought to defend the claim that a material structure can be endowed with thought and sensation through the internal forces of matter, and wished to cut off this path to (vital) materialism in favour of the traditional doctrine of the spirituality and immateriality of the soul (Euler 1842: 316-317).



strategies: whereas Euler argued that the mathematical concept of attraction could be completely covered by an impulsionist interpretation by invoking the aether hypothesis, Maupertuis argued against the possibility of excluding attraction a priori as a property of matter on the basis that it is contradictory to the most basic properties of matter:

I only intended to examine whether attraction, even when considered as an inherent property of matter, was metaphysically impossible. If it is such, even the most pressing phenomena couldn't make it accepted, but if she involves neither impossibility nor contradiction, we can examine whether the phenomena prove it or not at liberty. Attraction is no more, so to say, than a matter of fact; it is in the system of the universe that we must go look whether it is a principle that in fact occurs in nature, to what extent it is necessary in order to explain the phenomena, or finally whether it is introduced uselessly in order to explain facts that we explain well without it. (Maupertuis 1768a: 103-104; my translation)

What is remarkable about Maupertuis' account is, however, that it does not serve to separate metaphysics from physics, as is indicated by two peculiar features of this passage from the *Discours sur les différentes figures des astres*. First of all, the issue of attraction *can* be settled by metaphysics, *if* the notion can be proven to be a priori impossible; it is just the case that it doesn't settle it. Secondly, physics *could* reveal whether attraction is to be taken as a genuine property of matter. On the whole, then, I must concur with Lisa Downing's (2012: 298) assessment that "The *Discours*, in the end, veers closer to a genuine dynamicism or realism about attraction than at first appears. And, while it maintains that physics can *function* separately from metaphysics, it suggests that each may still have implications for the other."

It would be dishonest, however, to claim that these points are completely overlooked by the traditional interpretation. Indeed, Michael Friedman (1986: 28) attempted to refute the idea that Kant's disagreement with Newton is over matter theory. In his opinion, the "mathematical-mechanical" conception of matter which Kant rejects is not a Newtonian conception of matter, but an essentially Cartesian one. This assessment, however, overlooks the ambiguity present in Newton's matter theory itself. As Zvi Biener and Chris Smeenk (2012) have convincingly argued, Newton's account of matter in the (earlier editions of the) *Principia* combines, in a problematic way, the older "geometrical" conception of matter, i.e. in terms of extension, with the "dynamical" conception, i.e. in terms of force, for which the work is famous. This problem was pointed out to Newton by Roger Cotes, whose criticisms and conclusions show clear similarities to Kant's: Both Cotes and Kant believed Newton to be mistaken in his metaphysical hesitance towards gravitation as an intrinsic property of matter, and both believed that the step of Newton's argument for universal gravitation in Proposition VI of Book 3 essentially committed him to the dynamical conception (Cf. AA IV: 515). We

can therefore plausibly assume that Kant had a similar intuition about the ambiguity of Newton's conception of matter in mind.<sup>6</sup> Although Friedman (1992b: 156-157) acknowledges this strong parallel between Cotes' and Kant's point he points out that here, "Newton is criticized for not daring to be 'Newtonian' enough". What is peculiar about this assessment is that it criticizes Newton for not being more metaphysically daring, i.e. for adopting the anti-metaphysical humility that is characteristic of most early 18th century Newtonianism. As a result, we cannot have Kant both as an apologist of non-metaphysical, positive Newtonian Science *and* as the champion of a revolutionary Newton-inspired matter theory. More generally, if Kant is a Newtonian, he is so only in some respects; but then again, given the absence of a single unified Newtonian Science during his formative years, this should hardly come as a surprise.

---

<sup>6</sup> This criticism of Friedman of course fits well with the one voiced by Gordon Brittan (1986: 75-76). Brittan remarks that Kant's "mathematical-mechanical" conception of matter seems to be the one offered in Definition 1 of the *Principia* ("Quantity of matter is a measure of matter that arises from density and volume jointly" (Newton 1999: 403)), which does have a Cartesian overtone. Biener and Smeenk (2012: 114-115), of course, point out that 1) in his further reasoning, Newton uses instead the measure of matter offered in Definition 3 ("Inherent force of matter is the power of resisting by which every body, so far as it is able, perseveres in its state either of resting or of moving uniformly straight forward" (Newton 1999: 404) rather than that of Definition 1 throughout the *Principia*; and 2) in the clarification of definition 1 Newton ends up nearly equating the quantity of matter with the more dynamical measure of weight. Yet, there is an essential agreement over the fact that it is Newton's third *regula philosophandi*, and in particular the conclusion that impenetrability is a universal property of matter that reveals Newton's lingering commitment to the "geometrical" conception of matter. It is important to note here that Kant takes issue not only with Newton's listing impenetrability as an essential property of matter, but also with the problem of regarding inertia as a *force*, even though this forms the core of what Biener and Smeenk regard as Newton's *dynamical* conception of matter (Biener & Smeenk 2012). This problem has been noted by most commentators (e.g. Cohen 1999: 96), and has been regarded as a (perhaps unintended) analogue of the Leibnizian notion of *vis inertiae* (Garber 2009: 176-178). Moreover, it was already alerted by Euler, who sought to eliminate it from proper physical discourse because "inertia marks rather something completely opposed to the idea of forces" (Euler 1752: 423; my translation). It is important to note that Kant would, throughout most of his career, continue to regard inertia as a force, until he changed his mind in the *Metaphysische Anfängsgründe*, where he separated dynamics and mechanics. The characterization of forces, in casu attractive and repulsive force, are the proper domain of dynamics, whereas the determination of the laws of motion is that of mechanics. The repulsive force, which is meant to underlie an object's resistance to impressed force, is no longer the principle of inertia, but the metaphysical ground of *elasticity* (AA IV: 500). Inertia, on the other hand, is denied the status of force for the reason that it would imply an internal principle of change, something which violates the general principle of Kantian Mechanics (AA IV: 544). In this later work, then, Kant seems to have offered a reconciliation of Leibnizian and Eulerian positions by disentangling the notion of force from that of a law by assigning them to different subdisciplines of natural science.

### 1.1.2 Kant as Scientist

Perhaps most representative of the received view of Kant's own activities and credentials as a scientist is Erich Adickes massive two-volume study *Kant als Naturforscher* – representative because its title soon turns out to be a *contradictio in adjecto*, since it advances at the very beginning that Kant was *not* a scientist, but a philosopher who was ultimately a dilettant in scientific matters (Adickes 1924 I: 5). This idea was suggested, undoubtedly, by the *Marburger Schule* idea that Kant differed from major predecessors of his such as Descartes or Leibniz in not having made any real contributions to properly scientific progress. This idea runs into the expected *prima facie* objection that Kant wrote several works that seem to be intended as scientific papers, not reflections on scientific theory or practice. One of Adickes great contributions has been that he clearly indicated in what sense Kant differed, even in his patently “scientific” works, from the two aforementioned thinkers:

These two philosophers were in their full cast of mind real natural scientists and mathematicians as well. Not Kant. He never knew how to use the two important instruments because of which modern natural science became great: the experiment and mathematics. (Adickes 1924 I: 6; my translation)

Adickes goes on to indicate what we may regard as the standard view on Kant and his scientific writings:

The new great ideas that he contributed rested on intuitions and *aperçus*, with which he was endowed not as natural scientist in the strict sense, but as a generally richly talented mind and scientific genius. Intuitions and *aperçus*, however, cannot be commanded or regulated or compressed into a method; on their basis alone a continuous development of science will never be possible. That can only exist when they are prepared and completed by methodical investigation that further develops the problems organically; if they result, time and time again, in faithful, never exhausting careful work, in which it is imperative to experimentally validate, elaborate individually and when possible exactly calculate what was seen in general outlines in intuition. (Adickes 1924 I: 16; my translation)

Adickes' claim here seems to be that Kant was not scientifically minded because he consistently limited himself to the formulation of ideas on the basis of a general insight, but never cared to test these ideas experimentally and through exact measurement and calculation. He also claims that this latter enterprise is the proper scientific one, and marks it off from other types of human inquiry, something which he believes Kant himself fully realized. But when spelled out in this manner, it becomes possible to identify, in this smooth and rhetorically somewhat inflated orientation of Kant in the history of modern science, the two implausible assumptions on which it rests. On the

one hand, Adickes' claim that Descartes and Leibniz were representatives of the kind of science he is portraying here has lost much of its lustre over the course of the twentieth century. On the other, it seems to ascribe a practical contradiction to Kant.

The first assumption is the idea that Descartes and Leibniz were scientists in the way that figures such as Huygens or Hooker were, and that they relied on experiment and strict mathematical demonstration to advance their views. It is, I believe, a fair assessment that this image of early modern science has undergone crucial changes over the course of the century that divides us from Adickes' epoch. In this picture, there is an important distinction between the kind of scientific enterprises that philosophers like Descartes and Leibniz were pursuing, and the more Baconian-inspired style of science promoted by and through the royal society. In the *Principia Philosophiae* and the *Essay de Dynamique*, experiment and strict demonstration often give way to (at times strained) analogies, qualitative assessments and metaphysically grounded derivations. In Descartes's case, one need but think of his crucial error in the formulation of the law of uniform acceleration in free-fall<sup>7</sup>, the implausibility of his rules for the impact of moving bodies<sup>8</sup> and the bold conjecture that is his theory of vortices. In Leibniz' case, an analogous situation occurs with respect to his rather restricted use of the law of free-fall in Galilean kinematics, his largely qualitative account of perfectly elastic impact<sup>9</sup> and his unflinching use of harmonic proportions in cosmology (Cf. Westfall 1971: 306).

---

<sup>7</sup> Cf. Koyré (1966: 115) and additionally his lack of desire in continuing Galileo's enterprise, as expressed in his letter to Mersenne of October 11 1638 (AT VI: 380).

<sup>8</sup> There is some debate over the absence of a mathematical treatment of the laws of motion in Descartes. The best known position is the one advanced by Koyré (1966: 134-135), who maintains that there are two reasons for the absence of a mathematical physics in Descartes. The first one is that Descartes could not reason with idealized cases like Galileo could: given the fact that motion is always in a plenum and determined by the plenum, abstracting from the surrounding bodies does not give us an idealized case of movement, but an inconceivable case. The other bar to mathematization is that of "excessive geometrization", which led Descartes to overlook the mechanically relevant factor of time. At least the first reason is still usually appreciated (cf. Nelson 1996). Daniel Garber (2000:126) has proposed an alternative explanation by suggesting that the doctrine of the plenum makes it impossible for Descartes to draw conclusions on measures from empirical cases – that is to say: cartesian science does not allow for experiments like Galilean science does. In spite of the differences, however, both readings suggest that Cartesian science has trouble with at least one of the main traits ascribed to it by the Marburger Schule: mathematical treatment and experiment.

<sup>9</sup> The most consciously quantitative assessment of the rules of impact in Leibniz work is to be found in the manuscript *De corporum concursu*, a manuscript that has come under the broad attention of scholars only due to Michel Fichant's efforts. In this early work (from 1678), Leibniz develops his earliest conception of "vis viva" by contrasting and systematically assessing the estimations of impact according to the older Cartesian theory and the more recent account offered by Huygens, Mariotte, Wren and Wallis. In future developments, Leibniz would insist rather on the more general conceptual and metaphysical benefits of his account than on its empirical and mathematical virtues.

It is true, of course, that in contradistinction with Kant, both Descartes and Leibniz made considerable contributions to mathematics, and that the Marburger-Schule-reading relies heavily on this fact. Ernst Cassirer, for instance, said of Descartes that his entire foundations of mathematical physics derived from his introduction of analytical geometry (Cassirer 1922: 440-441), and regarded Leibniz to be the proper heir to this project, since he too derived his whole philosophy from his logical and mathematical inventions (Cassirer 1902: 102). In this way, he evokes the picture of two thinkers of idealistic temperament who were instrumental in providing mathematics as the core of proper scientific reason: their major contributions, analytic geometry and differential calculus respectively, constitute two of the most important moments in the process of mathematization (which is also one of arithmetization (Cf. Demarest 2013a)) and hence the autonomization of science as a positive enterprise from natural philosophy and metaphysics. But this reading is dissatisfying because Descartes's and Leibniz's mathematical innovations did not translate that obviously to their physical work. For instance, Descartes's introduction of analytical geometry did not stop him from failing to take the vectorial aspect of motion into account by leaving "determination" out of his "quantity of motion".<sup>10</sup> Similarly, Leibniz did not use his great contribution to mathematics, infinitesimal calculus, to quantify the perfectly elastic impacts that he invoked in his dynamics. As a result, the physics of Descartes and Leibniz is less intimately related to their mathematics than the Marburger reading suggests.<sup>11</sup>

The updated view appears clearly in Daniel Garber's distinction between Descartes and Leibniz on the one hand, and Galileo and Newton on the other:

---

<sup>10</sup> This fact has puzzled many a commentator, mostly since Descartes came very close to the formulation of the vectorial aspect of motion in his *Dioptrique*. The analysis of refraction offered there distinguishes between the force of impact and the *determination* of the direction of movement, a distinction that Abdelhamid Sabra has mapped onto that of scalar and vectorial quantity (Sabra 1981: 121). Interestingly, though, this distinction has not served to observe the two aspects of motion in force, but rather to prevent their proper integration; for although Descartes regarded them both as distinct aspects of the one primary mode of extension that is movement, he clearly regarded them as *modally distinct*, i.e. different secondary modes of movement (Cf. AT XI : 8-9 ; AT IV: 185). I tend to agree with Alan Gabbey's (1980:255-261) judgment that speed and determination are two distinct, though equally quantifiable modes of motion. Descartes's peculiar rules of impact then state that in impact, the incompatible modes of the two bodies must be changed, and that this change must be the least possible (AT IV: 185), so that the inversion of direction in the case of impact between bodies moving with the same quantity of motion (volume times speed (appropriately expressed, in modern notation, not by  $mv$ , but by  $V|v|$ )) is a change in determination only (and in one composite determination of total determination at that (as is clear from the treatments of Gabbey 1980 and McLaughlin 2000), not in speed (which never has a vectorial component). The reason for this seems to me that the determination of motion is determined by the *shape* of the moving body and the surrounding bodies, whereas its speed is determined by its *size*.

<sup>11</sup> Martial Gueroult (1967) already protested against the dissociation between the physical and the metaphysical in Leibniz' philosophy in 1934.

Descartes is working in a broadly Aristotelian tradition of natural philosophy. His aim is ultimately to give a view of the world that includes an account of the behavior of bodies as such, but grounded in an understanding of the true first causes: the nature of bodies, the causes of their motion, the way in which the laws that govern their behavior are grounded in the first cause, i.e. God. A different strand was the Galilean project. Galileo's project was within the domain of mixed mathematics, as it was called, a quantitative account of the world that favored mathematical description over an account of the ultimate first causes. I would claim that Leibniz is an inheritor of the natural philosophical tradition of Descartes, and Newton is an inheritor of the mathematical tradition that Galileo followed. (Garber 2009: 179)

Although Garber's account is not without its problems (it fails to make sense, for instance, of "how the Principia got its name", i.e. of how the Principia could have merited the name *Principia Mathematica Philosophiae Naturalis*, which implies that its author deliberately sought to situate it in the tradition that Garber takes to run counter to the tradition of which Newton was the culmination (Cf. Cunningham 1991)), it is indicative of the distance that separates the early 21st century picture of the 17th century from that advanced by Marburger-Schule neo-Kantians like Ernst Cassirer.

But besides this problematic assumption regarding Descartes' and Leibniz' engagements with science, a second, more vexing implausibility mars the Adickes-style assessment. In fact, it ascribes to Kant, whom it otherwise portrays as a philosophical champion of positive science, a scientific procedure that he himself would condemn as gruesomely misguided: a groping, analogical, unrigorous one. I would like to invoke a methodological principle of the scientific enterprise I am currently engaged in and suggest that such a blatant inconsistency should only be ascribed to so great a thinker if there is no satisfactory alternative explanation at our disposal. I will offer such an alternative interpretation in this chapter, arguing that Kant is more concerned with pointing out the limitations of early 18<sup>th</sup> century physics and making room for the kind of analogical reasoning that could emend these limitations.

### **1.1.3 The Looseness of Fit between Transcendental Philosophy and Natural Science**

The two premisses of the Marburger-reading that I criticized in the previous subsections both pertained Kant's place in the history of science. I have attempted to show, there, that the image of early modern science governing the Kant-literature has lost much of its plausibility in light of more recent studies of that particular enterprise. Nevertheless, none of this counts as an argument if it can be shown that, however wrong this view of science and its progress may be, it is in fact the one to which Kant

himself adhered. As evidence for this thesis, interpreters often invoke Kant's alleged bias towards physics – specifically, Newtonian mechanics. This bias, many believe, can be discerned in his cavalier attitude towards the other sciences (if they deserve this title of honor at all) and the close association between his transcendental principles and the Newtonian “Laws of Motion”. In this subsection, I will focus on the latter pillar, and reserve discussion of the former for the next subsection.

The locus classicus for any *Marburger*-style interpretation is the analysis of the analogies of experience from the *Critique of Pure Reason* in relation to the three fundamental laws of mechanics as they appear in the *Metaphysical Foundations of Natural Science*. This analogy is suggested by Kant himself, since he clearly states that the division of the *Metaphysische Anfängsgründe* is to follow the table of the categories closely (AA IV: 473-474). Nonetheless, I believe it is a mistake to overstretch this particular analogy, as Hermann Cohen, for instance, has famously done, in exclaiming triumphantly of the second analogy: “And wherein consists, then, the form of change? In nothing else than the succession of intensive magnitudes. In this succession, however, substance already acts, is force, namely inertia” (Cohen 1885: 463-464; my translation). In this seminal analysis, the foundations for Newtonian Mechanics allegedly provided by the *Metaphysical Foundations of Natural Science* is already the core of the argument from the first *Critique*.

This argument has lost most of its appeal over the almost 150 years since it first surfaced, although it still guides many contemporary readings of Kant. A first objection to it is that Kant's employment of the table of the categories as an organizing scheme for the discussion of the metaphysical principles of philosophy of nature is hardly a reason to believe these two tables are connected in a privileged manner. Indeed, Kant uses the same scheme throughout his entire Critical period to organize discussions of the most disparate topics, to the dismay of many interpreters, who believe Kant's architectonic usually obscures rather than reinforces his accounts. Taking Kant at his word in this one instance seems warranted only because of the preexisting bias towards Kant's philosophy of physical science, and the concomitant neglect of the other, more exotic occurrences of the table. This point is more vexing since the main reason Kant adduces for following the table of the categories in this instance is that, to him, it guarantees exhaustiveness (AA IV: 473-474). It is thus at least as plausible that in this instance as in all others, Kant, for better or for worse, relies on the *architectonic* features of the table of categories rather than on its specific content.

Besides this concern, a deeper concern saps the foundations of the *Marburger* rendition of the analogies, a problem that is known in the literature as the attribution to Kant of a “non-sequitur of numbing grossness”. This diagnosis of the failings of Kant's arguments was made twice, the first time by Arthur Lovejoy in 1906, where it appears in the following form:

If, says Kant, I am to regard a series of changes as constituting a perception of change, eine Begebenheit, I must ascribe to any such change in der Zeit eine gewisse bestimmte Stelle. This, however, I remark, is equally true of changes in representations that I do not attribute “to the object” at all. (Lovejoy 1967: 302-303)

And again, in almost similar wording, but strangely lacking reference to the fifty years older paper by Lovejoy<sup>12</sup>, in Strawson’s seminal *The Bounds of Sense*:

Kant is under the impression that he is dealing with a single application of a single notion of necessity. In fact, he not only shifts the *application* of the word “necessary”, but also changes its *sense*, substituting one type of necessity for another. It is conceptually necessary, given that what is observed is in fact a change from A to B, and that there is no such difference in the causal conditions of the perception of these two states as to introduce a differential time-lag into the perception of A, that the observer’s perceptions should have the order: perception of A, perception of B – and not the reverse order. But the necessity invoked in the conclusion of the argument is not a conceptual necessity at all; it is the causal necessity of the change occurring, given some antecedent state of affairs. It is a very curious contortion indeed whereby a conceptual necessity based on the fact of a change is equated with the causal necessity of that very change. (Strawson 1966: 137-138; original italics)

Strawson’s version of the argument has become the more famous one, mostly because it figures in a seminal attempt to introduce transcendental arguments, and hence involves a considerable amount of theoretical refinement. For this reason, it is also the position that is usually targeted by those who seek to salvage Kant’s response to Hume from the charge of non-sequitur, petitio principii, or even blatant contradiction. This non-sequitur is the idea that “the irreversibility of the order of our representations in apprehension might be taken as a criterion that indicates the presence of a cause” (Watkins 2010: 163).

A discussion of the problem of the second analogy goes, of course, well beyond the scope of this chapter, and even this book, so I will focus on one important response that has tied the question in with the issue of the relation between the categories and the laws of nature. This response is usually ascribed to Gerd Buchdahl, who assumed that there is a “looseness of fit” between the transcendental and the empirical levels of causality. Buchdahl (1969: 651-652) starts off by conceding the grounds for the Lovejoy-Strawson thesis, referring to an ambiguity in Kant’s wording that “seems to suggest that

---

<sup>12</sup> Henry Allison (2004: 254) is equally surprised by the seemingly independent similarities between Lovejoy’s and Strawson’s criticisms of the argument for the second analogy.



he believes himself to have shown nature to be in principle law-like, in a sense different from that so far ‘proved’; as though he was implying that he had shown every empirical sequence to be subject to some empirical causal law or other”, which would seem to imply that “Kant is actually claiming that the law of causality, so far proved only transcendently, may be employed *directly*, by simply resorting to experience, to indicate which empirical sequences are real” (Buchdahl 1969: 652; original italics). This is a clear analysis of the concerns behind the *non-sequitur*-charge, an analysis, furthermore, that shows how it bears upon the *Marburger* reading:

It is however easy to see why the mistaken interpretation of Kant should so spontaneously insinuate itself with many of his readers. For, in the case of Euclidean geometry, there is [...] a possible interpretation according to which the specific character (i.e. Euclidean) of the transcendental concept (extensive magnitude) is balanced by the character of that geometry. And again, in the case of the foundations of ‘pure natural science’ [...] it may be said that to the transcendental (‘metaphysical’) principles there *corresponds* the synthetic *a priori* character of the laws of Newtonian dynamics (laws of motion). But even if these two cases were accepted at their face-value (and I argue that they need not be), they would constitute no more than *applications* of the transcendental principles. (Buchdahl 1969: 660; original italics)

It is worth lingering on the major features of Buchdahl’s account. He claims that the charge of *non-sequitur* can only be made if one assumes Kant to be, in the second analogy, in the business of showing why nature is law-like. This assumption is often made due to 1) Kant’s momentary lapses into carelessness with terminology; and 2) the background belief that Kant wishes to ground the principles of Newtonian Science in the Critique. Against this, Buchdahl advances that Kant is not engaged in providing the basis for causal laws, let alone the Newtonian ones, in the Transcendental Analytic. The issue of laws only surfaces later in the text, in the Appendix to the Transcendental Analytic, where Kant discusses the conditions and functions of induction in scientific reasoning. This kind of reading has been dubbed, by Henry Allison, the “weak interpretation” of the second analogy, “since it insists that the argument shows (and is intended to show) merely that every event falls under the schema of causality rather than under particular causal laws<sup>13</sup>. In Allison’s opinion

---

<sup>13</sup> Allison contrasts the “weak interpretation” with a strong interpretation that seeks to uphold the link between the principle of causality and that of a causal law in the second analogy. One important such reading that seeks to counter the Lovejoy-Strawson thesis is Paul Guyer’s, an explicit target of Allison’s. Guyer claims that Strawson’s criticism begs the question against Kant, because it assumes something to hold that he reads Kant as arguing against: the fact that we can determine the subjective sequence of events independently from the objective sequence of events (Guyer 1987: 257). In this way, Guyer believes the second analogy can yield a

this approach has at least two significant advantages. First, by weakening the connection between the causal principle and the basic principles of Newtonian physics, it allows for the possibility of finding in the Second Analogy a transcendental argument that is independent of an appeal to the science of Kant's time. Second, it makes it easy to offer a concise and definitive counter to the familiar Lovejoy-Strawson objection that by moving from the determinacy or necessity of the sequence of perceptions in a single instance of event-perception to the existence of a general law covering all sequences of that type, Kant is guilty of a non sequitur of colossal proportions. Since on this reading there is no such inference, there is no non sequitur. (Allison 1996: 81)

Allison made this comment in the course of presenting his criticism of a recent capable criticism of the looseness of fit by Michael Friedman. In his frontal attack on the Buchdahl attempt to loosen the fit between Newtonian Science and Kantian Critique, Friedman makes the following plausible point: it makes little sense to conceive of anything as subsumed under the general principle of causality without thereby committing oneself to it being subsumed under a general causal law or uniformity (Friedman 1992a: 170), at least in so far as one is still committed to a nomological picture of causality. He goes on, however, to offer a theory that seems rather untenable as an interpretation of Kant. I will first offer my reasons for the latter diagnosis, which will also be my argument against this strategy in Friedman's revival of the *Marburger* story, and then comment briefly on the more plausible point raised by Friedman.

The following passage may serve as a summary of Friedman's thesis in his own words:

Particular causal laws, for Kant, have a peculiar kind of mixed status: They result from a combination of inductively observed regularities or uniformities with *the a priori* concept (and principle) of causality. Insofar as particular causal laws merely record observed regularities they are contingent and *a posteriori*; insofar as they subsume such regularities under the *a priori* principle of causality, however, they are necessary – and even, in a sense, *a priori*. (Friedman 1992a: 174)

As sensible as this may sound, the crucial passages on which Friedman builds his reading complicate matters substantially. The first passage he adduces as evidence is the following *Reflexion*:

Empirically, one can of course bring rules to the fore, but not laws; like *Kepler* in comparison to *Newton*; for to the latter [i.e. laws] belongs necessity. Therefore, they are known a priori. Nonetheless, people always assume that rules of nature are necessary, because that is why it is nature, and that they can be intuited a

---

stronger result than that expected by the Buchdahl-reading without falling prey to the ever looming *non-sequitur* charge.

priori; which is why people call them laws *anticipando*. The understanding is the ground of empirical laws, hence of an empirical necessity, where the ground of lawfulness can perhaps be intuited a priori e.g. the law of causality, but not the ground of the determinate law. All metaphysical principles of nature are merely grounds of lawfulness. (R. 5414; AA XVIII: 176; my translation)

In this Reflexion, Kant draws a distinction between what the empirical can yield, i.e. regularities, and what causal laws yield, i.e. necessity. Besides that, he points out that, contrary to what the charge of non-sequitur involves, no empirical laws of nature can be arrived at a priori. The problem is therefore that of seeing how empirical regularities can be elevated to the rank of laws of nature. Friedman tries to make sense of this by focusing on the specific example Kant uses: the contrast between Kepler's and Newton's accounts of the relations between the periods of the planetary orbits. Friedman (rightly) believes he can make further sense of the point Kant is trying to make by relating this Reflexion to some passages from the Opus Postumum, which repay quotation in full:

It is namely a peculiar appearance in the field of the sciences that there was a point in time when their progress seemed completed, when the ship lay anchored, and when there was nothing left for philosophy to do in a certain occupation. — Kepler's three analogies had completely yet merely empirically measured and mathematically described the phenomena of the circular motion of the planets, without however having the slightest idea of the *moving forces* that, together with their law, could be the causes belonging to them. Instead of the *aggregation* of motions of Kepler, which contained empirically collected rules, Newton made a principle of the system of moving forces from efficient causes. Unity. (AA XXII: 521; my translation)

The laws of motion were sufficiently exposed by *Keplers* three analogies. They were all mechanical. *Huyghens* had also [explained] the composite but derivative movement through centrifugal force and the forces that continuously drive it (*vis centrifuga et centripeta*) [...] but everything that was realized remained empiricism of the doctrine of motion and still lacked a general and proper principle, i.e. a concept of reason from which one can infer *a priori* a law of the determination of forces as one would infer from a cause to an effect, and Newton gave this explanation by calling the moving force *Attraction*. Because of this it became noticeable that this cause derived immediately from the bodies themselves and not from the communication of motion to other bodies and thus was not effected mechanically but purely dynamically. (AA XXII: 528-529; my translation)

What Friedman overlooks in these passages, in my opinion, is first and foremost the evidence it provides for the “weak interpretation”. For, as Kant makes clear, Newton's merit is to have provided not (primarily) an explanation of the planetary motions, but a *systematization*. But systematization is the proper business of reason, not of the understanding. In this way, Buchdahl's claim that the proper search for empirical laws

is constituted, in Kant's philosophy, by the regulative principles of reason rather than the constitutive principles of the understanding, is vindicated by the very passage Friedman adduces to discredit it. All the evidence is still in favour of the "weak interpretation", however troubling its philosophical status may itself be; for indeed, remitting Friedman's reading does nothing to answer the genuine philosophical concerns that lay behind it.

The most important problem with the "weak interpretation" of the second analogy is that it seems to render the link between causality and nomology, if not nonexistent, at least opaque. As Friedman understandably pointed out, it seems difficult to see what it would mean for a principle of causality *not* to bring along commitment to the existence of laws. The problem discerned in the passages above is, then, of course, the familiar one that no generalization, no empirical regularity, can yield the necessity and strict universality characteristic of laws. As Kant points out, regularities fail as bases for *prediction* (and, notoriously, for counterfactuals<sup>14</sup>), whereas laws are generally supposed to provide for this. But I believe that, beyond this, Friedman has erred (along with many interpreters) in reading back the debates on laws of nature in 20th century philosophy of science into Kant's text. The problem there is, however, the result of a generally nominalist and, quite frankly, Humean consensus<sup>15</sup>. As a result, many believe that the difference between a law and other kinds of universal propositions is that the former have a further property over and above the universality which yields the kind of necessity that allows for predictability and counterfactuality<sup>16</sup>. That this is Friedman's picture, is clear from the question he poses for (what he takes to be) Kant's account:

*How do the transcendental principles inject necessity into empirical laws of nature so as to secure them a more than merely inductive status? How do judgments that merely record observed regularities or uniformities become truly and "strictly" universal via the addition of the concept of causality? (Friedman 1992a: 175)*

"Injection", "Addition", the very metaphors Friedman invokes provide ample proof that logical positivism still lurks in the back of his mind, or that he is still marked by the fallout that followed its fateful collapse. As a result, he is blind to the more promising answer that glistens through the passages he brings under our attention: passages that suggest that laws and regularities are of a distinct logical *kind*. For what Kepler provided

---

<sup>14</sup> Cf. Chisholm 1946; Goodman 1947.

<sup>15</sup> That twentieth-century philosophy of science is thoroughly inspired by Hume, is hardly an outrageous or even original claim. This does not mean that there is nothing to be said about the possible differences between Hume and his logical-empiricist enthusiasts (Cf. Rosenberg 1993).

<sup>16</sup> The Humean thesis then states "that there is no difference between a law-like or nomological generalization and a universal truth drawn from exceptionless accidental regularities" (Rosenberg 1993: 75).

us were mathematical expressions for the relations between the behaviours of a class of objects, i.e. defined extensionally<sup>17</sup>. On the contrary, Newton showed us that the relations between the planetary orbits are the way they are because planets *have mass*, and because mass relates to other mass in a certain (empirically scrutable) manner: through attraction. The other major feature is that “planet” is a sortal, whereas “mass” is not (it is, quite trivially, a mass-term). Even though I cannot (and should not) expand in this reading here, not just because it goes beyond the scope of this work<sup>18</sup>, I would still like to advance the following hypothesis: for Kant, laws express categorical, synthetical relations between the intensions of two (preferably non-sortal) concepts.<sup>19</sup>

---

<sup>17</sup> Kant tellingly added a footnote to his reference to Kepler in Reflexion 5414: “*Planetæ omnes etc.*” the two words that sufficed for him were those that stood in for the extensional characterization.

<sup>18</sup> In 5.4.1, I will return briefly to this thesis, because it will allow me to make sense of what is otherwise a systematic conflation in Kant and his contemporaries, namely the conflation between a hierarchy of concepts and a hierarchy of laws. Philip Kitcher (1984: 191-194) rightly insists on the fact that a scientific theory, for Kant, is concerned with how the properties of things relate. However, he seems to take for granted that talk of properties (intensions) and talk of entities (extensions) is either interchangeable or at least straightforwardly related in Kant, a supposition that I of course wish to deny. I also dissent, as will become clear in chapter 5, from the ascription to Kant of a straightforwardly hierarchically taxonomical (i.e. Porphyrian) conception of systematic order. Lastly, I want to alert the reader to the fact that Kitcher seems ultimately to read in Kant a conception of a law of nature that builds on the Humean consensus of 20th century philosophy of science, in stating that “for Kant, the laws of nature are just the generalizations that would figure in the best unifying system in the limit of rational inquiry” (Kitcher 1996: 412). To Friedman’s interpretation, then, which characterizes laws as generalizations with a certain place in a system of knowledge (deduced from basic principles), Kitcher opposes one that characterizes them as generalizations with a place in a certain system (the best system). That both these interpretations are tailored to a Humean picture, becomes clear once we see that together, the two alleged criteria yield *David Lewis Conception of a Natural Law*: “a contingent generalization is a *law of nature* if and only if it appears as a theorem (or an axiom) in each of the true deductive systems that achieves a best combination of simplicity and strength” (Lewis 1973: 73; original italics) (mind that the sole difference of note in this definition, namely the insistence that it must figure in *all* best systems, is simply due to the fact that Lewis doubts, as he indicates in a footnote, that we can choose one system as *the* best).

<sup>19</sup> The account I ascribe to Kant here has obvious affinities to the one offered by Fred Dretske, who does not unambiguously endorse it, but does express its implications for analytic metaphysics: “These are inflationary times, and the cost of nominalism has just gone up” (Dretske 1977: 268). Luckily for Kant, his transcendental idealism may allow him to uncouple the epistemological and metaphysical costs and so avoid hyperinflation of his ontology. Additionally, I do not think Kant consciously offers this picture as an alternative to existing theories (as Dretske does), but rather assumes it in his response to Hume. Indeed, much of our problems in understanding the relation between Kant and Hume is that the former adhered to a form of intensionalism, whereas the latter was more of a nominalist (or at least had nominalist intuitions), and an extensionalist (for Hume, what we mistakenly believe to be general idea, is in fact a particular idea (an idea of a particular) joined with a linguistic sign (general term) that signals the warrant for unrestricted substitution of that idea with any other idea belonging to the same resemblance-class (a class of particulars between which a certain transitive, reflexive and symmetrical relation of resemblance holds) (Hume 2009: 20)). As a result, we look in Kant for a way of importing the nomic into the Humean “mosaic of local matters of particular fact” (Lewis 1986: viii), hindered by the fact that we can only meaningfully regard the natural qualities of which physics speaks

The proper business of science is, then, to discover these relations on the basis of the empirical regularities that we can establish.

#### 1.1.4 Kant and the Inexact Sciences

The final important pillar on which the Marburger edifice rests is the constation that Kant deems physics an exemplary science – in fact, we may even read him as stating that it is the only proper science. This physics-centrism is something that we commonly read back into our modern ancestors, but in Kant’s case there is of course textual motivation for it. He writes, after all, in the *Metaphysical Foundations*: “I assert, however, that in any special doctrine of nature there can be only as much proper science as there is mathematics therein.” (AA IV: 470); and this criterion allows him to dismiss various scientific enterprises as proper sciences, such as Chemistry<sup>20</sup> and of course, notoriously, psychology<sup>21</sup>. When he states that “what can be called proper science is only that whose certainty is apodictic; cognition that can contain mere empirical certainty is only knowledge improperly so-called” (AA IV: 468), he places only theoretical physics in the class of proper sciences, and relegates all the rest to the rabble of false pretenders. Such a picture seems to decide the issue of Kant’s predilection for physics quite conclusively. But here, as elsewhere, first impressions can be deceiving.

First of all, we should take into account the reasons Kant adduced for denying the status of proper science to so many academic enterprises pursued even in his own time. As becomes clear from the quoted passages, Kant holds proper science to 1) contain, or depend on, a *pure* part that furnishes its *a priori* principles; and 2) apply (apparently by grace of the pure foundation) *mathematical* methods to its object of study. As many authors have noted, this two-pronged criterion is extremely strong. Philip Kitcher (1996: 408) has interestingly castigated the criterion by pointing out that “consideration of the sciences as they were actually developed in the seventeenth and eighteenth

---

through their local instantiations. As Earman and Roberts (2005: 2) put it: “The Humean base may be characterized – very roughly – as the complete set of basic facts not offensive to those who are skeptical of non-logical, necessary connections in nature. The base includes particular facts about the existence of physical objects and their occurrent properties and spatiotemporal relations to one another. Excluded from the base are irreducibly general facts (e.g., that all electrons are negatively charged, that there exists at least one electron), and facts that involve laws of nature or other non-logical, natural modalities (e.g., facts about causal relations, counterfactuals, and irreducible dispositions)”. For Kant, particulars are deeply entangled through their physical properties, and an event constitutes some such entangled state.

<sup>20</sup> “chemistry can be nothing more than a systematic art or experimental doctrine, but never a proper science, because its principles are merely empirical, and allow of no *a priori* presentation in intuition.” (AA IV: 471)

<sup>21</sup> “Yet the empirical doctrine of the soul must remain even further from the rank of a properly so-called natural science than chemistry.” (AA IV: 471)

centuries, and, even more, as they have evolved since, reveals Newtonian mechanics as a rare exception". Behind this judgment lie two major considerations.

The first is that concerning the ideal standard of science. This surfaces from Kitcher's criticism of the position he ascribes to Kant, and which he believes Friedman to be complicit in, namely that which views

real sciences as organized bodies of knowledge with clearly-articulated concepts and general principles, mathematically formulated in axiomatic form. After the demise of Aristotelianism, a number of prominent seventeenth century figures, most notably Descartes, seem to have retained the idea of a system of the world, and to have endeavored to construct one based on sounder principles. Kant is heir to this tradition, and sees in Newton's *Principia* an approximation to the desired system. Thus he presents an ideal of scientific knowledge, which Friedman appears to accept, that demands complete conceptual clarification and removal of all residual conceptual and empirical problems before a scientific theory can be considered fully justified. (Kitcher 1996: 408)

Kitcher undoubtedly feels the urge to criticize this picture because he sees it as forming the backdrop to many 20th century theories of science, most notably the notoriously physics-centered seminal theories of Ernest Nagel and Carl Hempel. Such a picture can prompt two distinct responses from the defenders of sciences other than theoretical physics. The first option is, of course, to reject this standard (as Kitcher seems to do in his response to Friedman). This might take the form of an argument that the structure of physical explanation may be adequate for the objects of physics, but that other objects require other methodologies. Interestingly, such a line of reasoning was pursued by a school of neo-kantianism that rivaled with Cohen's *Marburger Schule*, namely the Southwestern School led by Wilhelm Windelband and Heinrich Rickert.<sup>22</sup> The other option is to point out that at least some other sciences have, albeit belatedly, adjusted themselves to the standard. The history of post-kantian science and philosophy is replete with examples of the latter strategy, perhaps most notably concerning

---

<sup>22</sup> Windelband (1924) basically agrees with the *Marburger Schule* on the idea that Kant's philosophy was fashioned after the mold of Newtonian Science. It adds, however, that, while some sciences have acquired, since Kant's time, a method analogous to that of Physics (most importantly chemistry), another evolution in recent science is more important: that of the emergence of history, and with it the *Geisteswissenschaften* as a methodologically well-founded scientific enterprise. History, however is not and cannot be a science based on laws and generalizations, like physics. It is more focused on understanding individual and particular wholes in their "value" (*Wert*). Windelband is, moreover, emphatic in the idea that this outlook requires us to put the *Critique of Judgment* on centre stage. This strategy has recently taken up again by Rudolf Makkreel (1990), who tries to extract a theory of the *Geisteswissenschaften* from the *Critique of Judgment*, but interestingly criticizes Windelband for overlooking the import of the third *Critique* and focusing instead on the *Critique of Practical Reason* (Makkreel 1990: 168).

chemistry, biology<sup>23</sup> and psychology<sup>24</sup>. Those taking the latter course then often mock Kant for his misplaced denial of the very possibility of something that was subsequently realized.

The second consideration lying behind Kitcher's comment is that Kant somehow disregards the scientific practice in his own time. The qualification "somehow" is in place because most authors go on to express some bewilderment over the fact that Kant seemed peculiarly interested in those enterprises that he seemingly derogated. In fact, the puzzle of Kant's knowledge of and interest in the practices of the non-physical sciences is as old as the *Marburger* interpretation itself, since it surfaces, uncomfortably late of course, in Hermann Cohen's *Kants Theorie der Erfahrung*:

Is mathematical physics actually the whole of natural science? May we adopt the perspective that takes *partem pro toto*, that takes the system of mechanical causality to indeed encompass the whole of nature?

[...]

Besides mathematical, there had traditionally been another kind of natural science, the pursuit of which not even the founders of dynamics Galilei and Newton had disdained, and in which Kant too had taken an interest in his old age as well as in his youth – and not just critical but also descriptive natural science, which Kant had distinguished from the Theory of Nature as Description of Nature. The transcendental conception of laws of nature is conditional upon a clear observation of the distinction between both these methods and aims of inquiry into nature. (Cohen 1885: 508-509; my translation)

The essential problem is this: if Kant does not regard the non-physical sciences as proper sciences, the methods of which are warranted by the transcendental deduction, then how can he still regard them as meaningful enterprises? The solution is usually sought in two directions: firstly, it seems that Kant's association of non-physical sciences with *reflective judgment* points in the direction that they are somehow important to systematization; secondly, they may be linked to that other meaningful human enterprise that seems to elude the bounds of sense: ethics and morality. The

---

<sup>23</sup> The Modern Evolutionary Synthesis is, in many ways, an attempt to 1) clear the ground conceptually for the life sciences by clarifying and unifying the basic concepts of (Darwinian) Evolutionary Theory and (Mendelian) Population Genetics; and 2) make them amenable to quantitative analysis through statistical means. In the earliest phase, of course, these two tasks went hand in hand (in the works of Ronald Aylmer Fisher (1918, 1930), John Burdon Sanderson Haldane (1924) and Sewall Wright (1932)). In the second phase, the mathematical aspect was less pursued, and attention shifted almost integrally to the conceptual core of evolutionary-genetical theory (Cf. Gould 2002).

<sup>24</sup> Gustav Fechner's idea of psychometrics was obviously inspired by the (neo-Kantian) idea that a proper psychological science should be able to quantify its object in a manner analogous to the quantification involved in measurement in physics (Fechner 1860: v).



latter option is more attractive to those who wish to embrace the idea that Kant is and remains a physics-centrist<sup>25</sup>, undoubtedly because it allows for a scheme for “compatibility” of contradictories. Nevertheless, I think taking Kant at his word requires us to focus more on the former option – and as I will try to make clear throughout this book, at the level of systematization the sharp boundaries between physics and the other sciences tend to blur. This is at least suggested by the Buchdahl-style “weak interpretation”, which holds that physical theorizing is a competence of reason, the capacity for systematization, not understanding, in its constitutive capacity. If so we may need to reverse the traditional question: we should attempt to think theoretical physics as a borderline case of Kantian science rather than as its normal form.

### 1.1.5 The Exception as Rule

In this Section, I have been engaged in undermining, one by one, the main pillars of the thesis that Kant’s philosophy is essentially linked to the fate of Newtonianism, or at least to (theoretical) physics. I have done so by offering a criticism of each of the major pillars of the *Marburger* interpretation.

The first pillar is the idea that Kant was taking account of the “Factum” of positive Newtonian science. I have responded to this by alerting the reader to the fact that, in Kant’s own time, it was not obvious that such a positive science had emerged as autonomous of metaphysics, and if it did, Kant’s enterprise seems to overstress the metaphysical problems<sup>26</sup>.

The second pillar is the central assumption that Kant differs from important predecessors like Descartes and Leibniz in his being a philosopher of science rather than a philosopher who practices science as well. I have tried to show that this assumption rests on a mistake about the relation between science and philosophy in Descartes and Leibniz, and a mistake about Kant’s own engagement with science.

The third pillar is the claim that Kant’s transcendental analytic is meant to directly provide the basis for Newtonian science in particular. I have responded to this by

---

<sup>25</sup> An important recent example of this is Alix Cohen (2009: xi-xii).

<sup>26</sup> It is important to note that my criticism differs from the one offered by Gordon Brittan (1978: 119), who states that “[o]ne of the difficulties for the claim that Kant’s project is to provide metaphysical foundations for Newtonian science, and thereby prove its validity, is that, to my knowledge, Kant himself never characterizes it in that way. Far from guaranteeing physics from skeptical attack, the task is to say how metaphysics can become, like physics, a science”. I do not take for granted the idea that, for Kant, physics does not stand in need of defense.

defending Buchdahl's thesis that there is an irreducible "looseness of fit" between the principles of the understanding and the laws of nature.

The fourth pillar is the claim that Kant clearly preferred physics over other sciences. I have challenged this by showing that Kant was clearly engaged in other sciences, and as such had to provide some legitimation for them, albeit a different kind of legitimation than we would initially expect.

All of this is an essential *via negativa* intended to clear the way for my revised picture, which, I believe, in part draws its strength from the fact that it responds to the internal problems of settled interpretations. Moreover, I believe it is necessary to combat the entrenched picture we have of science, a picture that largely derives from late 19th century narratives, and that describes the gradual emergence of a positive scientific enterprise. For all its drawbacks, this picture is still read into Kant, even if it is often immediately conceded that he was mistaken in this. As opposed to this, I believe Kant was more responsive to two other features of physical science, namely 1) its relation towards, and ultimate embedment in, the tradition of natural philosophy; and 2) its usage of conceptual means in getting at results. This is suggested by a key passage from the preface to the second edition of the *Kritik der reinen Vernunft*, where Kant writes the following:

When Galileo rolled balls of a weight chosen by himself down an inclined plane, or when Torricelli made the air bear a weight that he had previously thought to be equal to that of a known column of water, or when in a later time Stahl changed metals into calx and then changed the latter back into metal by first removing something and then putting it back again, a light dawned on all those who studies nature. They comprehended that reason has insight only into what it itself produces according to its own design; that it must take the lead with principles for its judgments according to its constant laws and compel nature to answer its questions, rather than letting nature guide its movements by keeping reason, as it were, in leading-strings; for otherwise accidental observations, made according to no previously designed plan, can never connect up into a necessary law, which is yet what reason seeks and requires. Reason, in order to be taught by nature, must approach nature with its principles in one hand, according to which alone the agreement among appearances can count as laws, and, in the other hand, the experiments thought out in accordance with these principles - yet in order to be instructed by nature not like a pupil, who has recited to him whatever the teacher wants to say, but like an appointed judge who compels witnesses to answer the questions he puts to them. Thus even physics owes the advantageous revolution in its way of thinking to the inspiration that what reason would not be able to know of itself and has to learn from nature, it has to seek in the latter (though not merely ascribe to it) in accordance with what reason itself puts into nature. This is how natural science was first brought to the secure course of a science after groping about for so many centuries. (CPR B xiv)

The Science with which we find Kant engaged here is not the positive science alleged to have developed in the wake of Newton’s groundbreaking *Principia*, but the natural philosophy that constituted an ongoing project from the late 16th to the mid-18th century. Instead of Newton, we find much earlier theorists like Galileo and Toricelli, and instead of a “respectable” recent mathematical physicist like d’Alembert, we find a mention of Georg Ernst Stahl, a radically “vitalist”<sup>27</sup> opponent of the mechanical approach in natural history. Finally, what Kant focuses on in these passages is the amount of background assumptions researchers consciously make in order to let phenomena be meaningful and instructive. In physics, interesting results come up only because of crucial abstractions and conscious conceptual decisions, in experimentally highly contrived circumstances. It is the exceptional role of physics in staging exceptions that drew Kant’s attention – and it is the harrowing backlash this has for the understanding of all that resists such practice by which he is touched. Specifically, I will argue in the next two sections that Kant took issue with the radical determinism of both Leibnizian and Newtonian mechanics, which he believed to be contradicted by the specifically dynamic nature of some systems. In the next chapters, we will see how this is relevant to his views on the temporal evolution of biological systems, human nature and even our cognitive apparatus.

## 1.2 Finding the Dynamical in Dynamics

### 1.2.1 Sufficiency and Continuity in Leibnizian Dynamics

In order to understand Kant’s earliest engagements with physics, we need to first take a closer look at Leibniz’s dynamics. This development was a major topic of discussion among continental European natural philosophers, both for its scientific and its philosophical implications. One major question here was whether the basic assumptions of the new science of dynamics squared with those of the previously dominant mechanical outlook and, if not, which was preferable. Another was whether the assumptions of the Leibnizian and the Newtonian versions of mechanics squared with

---

<sup>27</sup> My use of scare quotes here is meant to indicate my general unease with this term. As Charles Wolfe (2011) has shown extensively, not only is it marred by the connotations it acquired over the course of the 20<sup>th</sup> century, but it is also a much more versatile and diverse position than is traditionally suggested. I still use the term here, however, because even Wolfe seems to regard Stahl as a rather typical representative of the position most widely understood under it.

each other and, if not, which was preferable. As we will see, Kant preferred dynamics to mechanics, and incorporated both recognizably Leibnizian, i.e. anti-Newtonian, and Newtonian, i.e. anti-Leibnizian, elements in his own account.

Leibniz greatest contributions to physics are usually regarded to be his criticisms of the (according to some moribund) Cartesian mechanics of the likes of Malebranche, Cordemoy, de la Forge and Régis. With enormous penetration, he saw through the fatal flaws of Descartes's conception of motion and was able to synthesize an underlying conceptual, even metaphysical disagreement out of what was already, due to the efforts of Christiaan Huygens (1669), Edme Mariotte (1673), Christopher Wren (1668) and John Wallis (1668), a scientific challenge to the reigning orthodoxy. In presenting this criticism of Descartes, he relied on substantial philosophical commitments. As a result, for the educated 18th century German readership, the new science of "dynamics" came to be associated as much with these underlying commitments as with its empirical upshot.

The first challenge Leibniz raised to the Cartesians is by far the most famous and most highly publicized one, because it resulted in the infamous *vis viva* dispute that would rekindle time and again amongst continental physicists, even after D'Alembert provided his triumphant dissolution of the difficulty in his *Traité de Dynamique*<sup>28</sup>. For the purposes of the present argument, it is not necessary to plough through the morass of vague or confused concepts that characterized this debate, mostly because Leibniz himself seems to have mostly been concerned with the metaphysical and methodological lessons to be drawn from his demonstration of the errors of the Cartesians (cf. Iltis 1971: 26-27).

Although the evidence suggests that Leibniz had formed the idea as early as 1678<sup>29</sup>, it is only in the correspondence with Arnauld and the infamous paper "Brevis demonstratio erroris Cartesii et aliorum circa legem naturalem, secundum quam volunt a deo eandem semper quantitatem motus conservari, qua et in re mechanica abutuntur", published in the *Acta Eruditorum* in 1686, that Leibniz first voiced his criticism of the Cartesian measure of force. In that paper, Leibniz contrasts his own *vis motrix* (motive force), measured through the height to which it can elevate a body (GM VI: 118), with the Cartesian quantity of motion, which is measured as  $V|v|$ <sup>30</sup>. According to

---

<sup>28</sup> The case of this book is interesting in itself because of the self-professed inadequacy of its title; for D'Alembert knows full well that the term was coined by Leibniz as "Science of forces and motive causes" (D'Alembert 1758: xxxi; my translation), whereas D'Alembert's solution to the *vis viva* dispute consisted precisely in the claim that we should focus on *effects* rather than on *causes* in physics. Hence, D'Alembert denied the possibility or desirability of a proper science of dynamics beyond the study of motion (mechanics).

<sup>29</sup> See footnote 9 of this chapter.

<sup>30</sup> I use this formula to avoid the confusion, common to late 17th and early 18th century science, of what we would regard as momentum ( $mv$ ) and Cartesian Quantity of Motion. See also footnote 10 of this chapter.

Leibniz, the metaphysical import of this is that forces are to be considered as basic to nature. He argues for this in two distinct ways.

The first strategy is the better known of the two, at least among students of early modern science and philosophy. It is patently not Leibniz's primary intention to offer this criticism of Descartes in order to show how we should go about measuring mechanical interactions, but rather to show the insufficiency of the Cartesian picture of nature. Descartes famously argued that 1) nature is to be understood in terms of extension and its modes alone<sup>31</sup> (AT VIII: 42); 2) the universe is subject to a conservation law, stating that God always preserves the same quantity of motion (and rest) in the universe (AT VIII: 61); and 3) motion is a mode of extension (AT VIII: 25). Taken together, these imply that we can grasp the world solely through its geometrical properties. Leibniz combats this picture by showing that it leads to trouble for the second assumption. While he agrees with Descartes that the universe should be subject to a conservation law, and that from this it follows that in every mechanical interaction something must be conserved completely, he denies that what is conserved is quantity of motion.

He argues for this in the following manner: take two objects, one of 1 unit of weight falling at the rate of 4 units of speed, and another of 4 units of weight falling at the rate of one unit of speed. According to the Cartesian measurement, these two bodies will fall with the same "force" (i.e. quantity of motion. Because of Galileo's principle of the lever ( $mv^2$ ), however, the force generated in this way would be able to elevate the body of 4 units of weight to an altitude of 1 unit of distance, whereas it would elevate the body of 1 unit of weight to an altitude of a whopping 16 units of distance. We could thus in principle construct a machine that generates a given amount of force by dropping a big object, uses the generated force to lift a small object, and then drops the small object in order to generate a much greater force. Such a machine could in principle generate an unrestricted amount of force using only a restricted initial force, and would constitute a perpetuum mobile. But as Leibniz notes, the idea of a perpetuum mobile is an absurdity, and Descartes will admit this as well (GM VI: 124-125). For Leibniz, the conclusion is absurd because it violates one of the many versions of his infamous *principle of sufficient reason*, namely that of the *equality of cause and effect*.<sup>32</sup> Leibniz seems to conclude from

---

<sup>31</sup> Some qualification of this thesis is required, as has been shown by Daniel Garber (1992: 66-67), for Descartes's body is of course subject to a set of attribute-transcending "general notions", such as, importantly, duration. Without the latter notion, Cartesian science would not get off the ground, and yet it is not clearly a mode of extension. However, this does not seem to complicate things for the argument I am trying to develop here.

<sup>32</sup> This is one of the earliest incarnations of the principle of sufficient reason, first formulated in the 1676 manuscript *De Arcanis motus et Mechanica ad puram Geometriam reducenda* (cf. Hess 1978: 203-204).

this that the proper metaphysical substrate of physical reality cannot be geometrical extension, but is rather dynamical force.

A decade later, in the *Specimen Dynamicum*, Leibniz offered a second line of argument in favour of his dynamics and against the merely mechanical outlook of the Cartesians:

Strictly speaking, motion (and likewise time) never really exists, since the whole never exists, inasmuch as it lacks coexistent parts. And furthermore, there is nothing real in motion but a momentary something which must consist in a force striving toward change. (GM VI: 235; Leibniz 1989: 118)

In order to understand what Leibniz is arguing here, it is important to note that the *Specimen Dynamicum* is the text where he introduces his distinction between living force and dead force:

One force is elementary, which I also call *dead force*, since motion does not yet exist in it, but only a solicitation to motion, as with [...] a stone in a sling while it is still being held in by a rope. The other force is ordinary force, joined with actual motion, which I call *living force* [...] which arises from an infinity of continual impressions of dead force. (GM VI: 235; Leibniz 1989: 121-122)

If we read these two passages in tandem, and relate Leibniz's argument against Descartes's geometrical conception of nature to his distinction between dead and living force, the following picture arises: motion cannot be conceived as actualized in a single instance of time, since it is defined as translation of place over time. But if it is to be considered as real, it must somehow have parts which themselves allow for the specific time-relation that it requires. Leibniz believes the properties of velocity and direction, both of which are essential to motion, can only be instantiated at an instant in the form of a somehow finalistic force. In the unpublished *Essay de Dynamique* of 1692, the analogy is even clearer, since Leibniz surmises there that dead and living force relate to each other as point to line (Costabel 1973: 127). I read Leibniz's point here as follows: just as no single point can have length or direction, so too no single instant of motion can have velocity or direction. But points needn't be regarded as non-entities; for calculus allows us to regard them as infinitely small rather than as unextended. To Leibniz, the idea of an infinitesimal seems only to make sense in a force ontology, where the infinitesimal parts of a force can themselves have the properties required of them, i.e. both a scalar and a vectorial quantity. But, importantly, this means that a motion is already determined before it is actualized, namely by the solicitation to motion.<sup>33</sup> Here again,

---

<sup>33</sup> As evidence for ascribing this view to Leibniz, Robert Sleigh (1990: 118) has adduced the following passage from a 1687 letter to Bayle: "I would like to add a remark of consequence for metaphysics. I have shown that force ought not to be estimated by the product of speed and size, but by the future effect. However, it seems

the principle of sufficient reason rears its head, this time in the form of the relation between dead force and living force – for whence come direction and velocity if they are not already contained somehow in the first instance of motion? How can they emerge simply through the passing of time? As I will argue in the next section, these were the questions Kant would ask.

Before passing on to Kant's discussion of Dynamics, however, we must first take a closer look at Leibniz's challenge to Descartes's rules of impact, which stated that when two bodies with equal quantity of motion collide, they both rebound (reverse their direction) with their speed before impact, but that when bodies with unequal quantities of motion collide, only the one with the smaller quantity of motion is reflected. Many readers had already supposed that there was a problem with this account, and several authors had provided alternative solutions, most importantly Huygens, Mariotte, Wren and Wallis. But contrary to these authors, who had a more mathematical and less metaphysical agenda, Leibniz pushed a more general point, as he did in this passage from a letter to Bayle:

This difference of events in the two cases is not reasonable; for the inequality of the two bodies can be as small as you would like, and the difference between the suppositions in the two cases, that is to say the difference between such an inequality and a perfect equality could be smaller than any given [quantity]. In that case, in virtue of our principle, the difference between the results or events should be smaller than any given [quantity]. However, if the second rule were as true as the first, the contrary would occur, since according to the former rule, an increase, however small, of the body B that was previously equal to C, produces an

---

that force or power is something real at present, while the future effect is not. From which it follows *that we must admit in bodies something different from size and speed, at least unless one wants to refuse bodies all power of acting*" (G III: 48; Sleight 1990: 118). Daniel Garber subsequently adduced a further passage that clarifies Leibniz's point: "The notion of force is as clear as that of action and passion, because it is that from which action follows when nothing prevents it...On the other hand, motion is a successive thing, which, consequently, never exists, any more than time does, since all of its parts never exist together. Unlike that, I say, force or effort exists completely at each moment, and must be something true and real. And since nature takes account of that which is true rather than that which exists only entirely in our mind, one finds (as I have demonstrated) that it is also the same quantity of force, and not the same quantity of motion (as Descartes thought) that is conserved in nature. And it is from this principle alone that I draw everything that experience has taught about motion and about the impact of bodies, against Descartes's rules, and that I have established a new science which I call dynamics, whose elements I have set out." (Garber 2009: 154-155). The point of these passages seems to me that since motion is never given fully at an instant, neither can its characteristic features (speed and direction), unless these are somehow contained in something that can exist at an instant. Force, to Leibniz, is precisely something that can exist fully at a given point in time whilst still having properties that are only exhibited over a period of time. The consequence of this I believe Kant to have been worried about is that the eventual motion of a body is already contained in its initial state, since it is contained in the force fully existing at the initial state.

extremely great difference in effect, since it alters the absolute reflection into absolute continuation, which is a leap from one extreme to the other. (G III: 53; my translation)

What Leibniz is trying to convey here is that Descartes's rules of impact violate a basic principle of metaphysics, and are therefore certainly false. This principle is the law of continuity, which he consistently formulates as "if the arguments are ordered then the values are ordered as well" (G III: 52; GM VI: 250). Ultimately, this principle reflects Leibniz demand (not just belief) that "whenever one physical quantity or state is a function of another, that function is continuous [...] and everywhere differentiable" (Mates 1986: 166). There is, of course, a way in which it reflects the principle of sufficient reason as well, because it demands that every difference in the value is warranted by an equivalent difference in the argument.

Another consequence Leibniz draws from the law of continuity is just as important, for he believes it to prove conclusively that all situations of impact should be regarded as cases of *elastic collision*. As he puts it in the second part of the *Specimen Dynamicum*:

From this follows something Descartes opposed in his letters, something many gentlemen of great reputation are even now unwilling to admit, that *all rebound arises from elasticity*, which explains many elegant experiments that show that *a body is deformed before it is impelled*, as Mariotte nicely demonstrated. And finally, a most wonderful conclusion follows from this, that no body is so small that it is without elasticity, and furthermore, each body is permeated by a fluid even subtler than it is. And thus, *there are no elements of bodies*, nor is there maximally fluid matter, nor are there little solid globes (unintelligible to me) of the second element, both determinate in shape and hard. Rather, the analysis proceeds to infinity. (GM VI: 249; Leibniz 1989: 132-133)

This passage reveals a major reason for Leibniz's resistance towards the idea of (physical) atoms: he feels that the assumption of atoms, and hence of the occurrence in nature of cases of perfectly hard collision, would dislocate the causal links in reality. For, in the case of perfectly hard collision, the state at the instant just before impact and the state at the instant just after impact are radically different. Such an assumption, Leibniz fears, makes the causal link between the two instances unintelligible. Of course, he may very well have been right in this, because it is a reflection on perfectly hard collision that prompted Hume's infamous skeptical arguments regarding causation: "Motion in one body is regarded upon impulse as the cause of motion in another. When we consider these objects with the utmost attention, we find only that the one body approaches the other; and that the motion of it precedes that of the other, *but without any sensible interval*" (Hume 2009: 54; italics added). Hume thus assumes that causality is a relation that holds between two successive events and that their succession is discontinuous. One can respond to this by either denying that causality is a relation



between two events rather than a feature of a single (causal) event<sup>34</sup>, or by conceding that it is a relation between two successive events, but counter that the succession is perfectly continuous. Leibniz, and all those following him in his predilection for elasticity (such as Kant<sup>35</sup>) may very well have had, if not an explicit appreciation, then at least an important inkling of the looming threat of this conundrum.

In order to appreciate just how important these two metaphysical reflections on the foundations of Cartesian physics are, one need but point out that they function as the core of the discussions of the first and second parts of the *Specimen Dynamicum* respectively. This shows to what extent Leibnizian Dynamics was imbued with the metaphysical principle of sufficient reason. But it is also this principle which seems to bind the universe together in the strictest of determinations, and to make temporal development a mere formality after the done deal of the initial condition. As we will see in the remainder of this chapter, Kant would continue to appreciate the principle of sufficient reason for its efficacy in ensuring the causal and intelligible structure of the world, but remain equally vexed by the fact that it makes this causal and intelligible link between all things too strong.

### 1.2.2 The True Estimation of Leibnizian Dynamics

The first work Kant sought to submit to the reading public was, not coincidentally, an essay on the vis viva debate: the *Thoughts on the true estimation of living forces and assessment of the demonstrations that Leibniz and other scholars of mechanics have made use of in this controversial subject, together with some prefatory considerations pertaining to the force of bodies in general* from 1749<sup>36</sup>. In the eyes of many commentators, however, it was a premature publication, riddled with misconceptions and hopelessly outdated when it finally hit the press. As Martin Schönfeld put it:

Kant's philosophical debut was a false start. He later considered the *True Estimation of Living Forces* a thorough embarrassment, which, for all practical purposes, it was. Not only was Kant incapable of resolving the problem of force, but also unbeknownst to him, Jean Le Rond d'Alembert had already published a theory that effectively settled the debate three years before Kant turned his mind to it. (Schönfeld 2000: 18)

---

<sup>34</sup> This position is traditionally associated with Aristotle, and has been taken up again more recently by Nancy Cartwright (1994). Eric Watkins (2005: 398-399) ascribes this position to Kant, but, as my treatment should make clear, I do not agree with him in this.

<sup>35</sup> See footnote 6 of this chapter.

<sup>36</sup> Although the title page suggests that it was printed in 1746, the book was not published until 1749.

Although this judgement is largely correct, it does call for some caveats. First of all, Schönfeld suggests that we can see just how little control Kant had of the technical aspects of the debate if we appreciate that it had already been solved, by D'Alembert in the first edition of his *Traité de Dynamique* from 1743, and by Roger Boscovich in his *De Viribus Vivis* from 1745. Since these books contained a solution that required a serious technical grasp in order to be understood, Schönfeld surmises, the young and insufficiently educated Kant could not have picked up on them. There are difficulties with this argument, however. The first is that it seems far from obvious that D'Alembert had already solved the issue in the first edition of the *Traité de Dynamique*. Carolyn Iltis (1970) has suggested that only in an addendum to the second edition of 1758 can we find a proper response to the problem. Secondly, however, both D'Alembert (Hankins: 284) and Boscovich (1745: 30) expressed a preference for the estimation in terms of the (revised) *Cartesian* measure, which means that they can both be read as embracing one side of the debate. Finally and most importantly, the reservations both authors had towards *vires vivae* is that the latter are only conserved in elastic collision. For D'Alembert, then, the assumption that nature consists of perfectly hard solid atoms played an important role in his resistance towards *vires vivae*. Boscovich, on the other hand, would ultimately go on to provide an important template for the Leibnizian image of nature as exhibiting only elastic collisions through his idea of point-masses. It is well known that Kant would later develop a more than superficial sympathy for that view. In sum, I believe we should not decide the issue over *vis viva* as an outdated one at this point, and keep in mind that important philosophical issues remained even in the wake of D'Alembert's and Boscovich's treatments.

Secondly, Schönfeld's judgement evokes a picture of Kant as a pitiful figure: a "technical philosopher" out of control of his technicalities, taken in by a fallacious proof of an absurd conclusion. This image may seem to keep us from discerning the major points Kant is trying to make here, and read into it instead the classic concerns that we are sure must play in, which is exactly what Schönfeld goes on to do:

Nonetheless, the *Living Forces* is fascinating. It reveals how the mind of the budding philosopher worked. Echoes of thoughts that had been formulated here first reverberate through the whole critical period, despite Kant's quick rejection of the treatise. Attitudes emerged here that were later transformed into the dominant motives of his philosophizing. Assumptions that Kant boldly introduced in the *Living Forces* later returned as problems requiring solution or claims needing explicitation, and as a result, many themes of the *Living Forces*—the beauty and perfection of nature, the tension between physical influx and preestablished harmony, the concepts of substance and world, the idea that force generates space—blossomed into the topics of the major precritical treatises in the next decade, the *Universal Natural History* (1755), the *New Elucidation* (1755), and the *Physical Monadology* (1756). (Schönfeld 2000: 38)

Later on he interestingly adds that “[b]ecause the *Living Forces* were in need of further clarification, the book became for Kant a list of things to do. Sections I and III generated the items on this list” (Schönfeld 2000: 54). There is an obvious reason why authors choose to focus on these two sections: they are more philosophical and conceptual in nature, distinct from the (misplaced) technicalities of the second section. But marred as the latter section is, it repays a closer look, especially the passages which Kant initially intended as a conclusion to the section. There he clearly alerts the reader to the general aim of the work: “Hence we do not have any dynamic principles at present from which we could justifiably proceed. Our work, which promises to present the true estimation of living forces, should make amends for this defect.” (AA I: 117) Kant thus seems clearly concerned with the basis for a dynamics, not with the proper measurement of force. This suggests indeed that, despite his enormous route through technical debates, he may ultimately have had mainly philosophical concerns here. I believe we can find out which these philosophical concerns are from his comments on Leibniz’s mathematical estimates of force. Indeed, Kant seems to push the point that the conception of nature in terms of *vis viva* violates precisely the two basic principles of Leibnizian science: the principle of equality of cause and effect (AA I: 104) and the law of continuity (AA I: 105). He therefore suggests that it is through the *principle of the best* (yet another manifestation of the principle of sufficient reason) alone that Leibniz can harmonize his estimation of force with his own fundamental principles. But this, Kant believes, is not the way to go:

But even this minor defense is flimsy. We are speaking only of the mathematical estimation of forces and it is no surprise if it does not match God’s wisdom perfectly. Mathematics is a science isolated from the medium of genuine knowledge, it does not sufficiently meet the rules of decorum and appropriateness if taken alone, and it must be combined with the tenets of metaphysics if it is to be perfectly applied to nature. The harmony present among truths is like the agreement found in a painting. If one takes one specific part away, then decorum, beauty, and design will disappear; all parts have rather to be seen together in order to perceive these same features. The Cartesian estimation is contrary to the designs of nature; it is accordingly not the true estimation of forces in nature, but this does not prevent it from being the true and justified measure of force in mathematics. For the mathematical concepts of the properties and forces of bodies are quite different from the concepts encountered in nature, and it is enough to have seen that the Cartesian estimation is not contrary to mathematical concepts. But in order to determine the true estimation of force in nature, we must connect the laws of metaphysics with the rules of mathematics; doing so will fill in the gap and better meet the designs of God’s wisdom. (AA I: 108)

The following reading of Kant’s intentions in the *True Estimation* suggests itself through these passages: Kant focused on an unfortunate tension between two major aspects of

Leibnizian Dynamics: the role the principle of sufficient reason played in it on the one hand, and the temporal evolution that it was supposed to ground on the other. As I tried to indicate in the previous section, Leibniz had two motives for asserting the reality of underlying forces: the fact that they and they alone allow for a measure that does not violate the basic principles of intelligibility (and, for Leibniz at least, ipso facto possibility), and the fact that they allowed us to understand how a state at an instant can have properties that only manifest over time (typically, velocity and direction). Kant suggests that there is a tension between these two arguments: the first seems to understand the initial state as a full determination of the consequent states, the latter seems to conceive of an initial state as something that is but one, underdetermining moment in a time series. Such a reading of Kant's argument is further strengthened by the fact that it explains the peculiarities of his theory of vivification.

In the third section of the *True Estimation*, Kant develops a theory of vivification which states that an object starts off with a *vis mortua*, and then gradually acquires a *vis viva* through the finite process of vivification:

This also implies, by the law of continuity, that the same body that possesses dead force in an initial moment and acquires living force in the next, a force that is to the former like an area is to the generating line, gains this force only in a finite time interval. For suppose we posited that it acquired this latter force not in a finite time interval after the initial moment, but instead instantly, in the infinitely short period after the initial moment, then this would be like saying that it already had this living force in the initial moment itself. For the law of continuity, and even mathematics as such, demonstrate that it does not make any difference whether I say that the body happens to be in the initial moment of its motion, or in the infinitely short period following it. But in the initial moment of motion itself the force is dead, and so we cannot say without contradiction that the force is therefore living if we also stated that this living force can be encountered in motion only after a finite interval, after the action of the external cause.

The body's natural force actually maintains within itself the externally received impression and since through its continuous striving it accumulates in itself the formerly point-like intension until it becomes like a line, which is proportional to the velocity-like force caused in it from without, it accumulates, on its own, the force obtained from the outside, which was previously only like a line too, until it is eventually like a plane whose one side represents the externally imparted velocity and force, while the other side models the intension, which is proportional to this externally imparted velocity and force, which has grown, on its own, from the body's interior. (AA I: 145-146)

Thus, according to the Leibnizian picture, there is a distinction between the dead force as it operates in the initial state of the system, and the living force as it appears after a finite time. But since the force is not already "alive" during the initial stage, it only becomes alive, vivifies, through a period in which an external force acts upon the body.

Kant describes this process of vivification as one where the internal force preserves, integrates and accumulates the externally received impressions of force. The difference here seems to be that, whereas for Leibniz the internal force is important because it already “contains” the later developments, for Kant it is important because it can come to internalize external influences over time and thereby acquire a force that is not reducible to the mere communication of external mechanical forces. Only over time can we speak of a proper living force, and only because of this time-factor do we need to take recourse to the metaphysical. In the realm of the mathematical, the principle of sufficient reason and its radical determinacy reigns supreme, and only in the natural, properly grasped through metaphysics, can time be a productive factor. That this is the view of the young Kant is clear from the second paragraph of the third section:

Mathematics does not permit its body to have a force unless it is wholly produced by the external cause of its motion. Accordingly, mathematics admits force in the body only insofar as force was caused in it from the outside, and hence one will *always find its force to the same degree in the causes of its motion*. This is a basic law of mechanics, whose presupposition, however, does not admit any estimation other than the Cartesian. But, as we shall soon show, the body in nature is of an altogether different constitution. *That body has the capacity to increase, by itself and in itself, the force awakened externally by the cause of its motion, which means there can be units of force in it that did not originate from the external cause of motion*, that may be larger than this cause, that therefore cannot be measured with the same yardstick as the one used for Cartesian force, and that accordingly involve another estimation. (AA I: 140)

Kant argues that the old mechanical perspective of Descartes, which regards all forces as externally caused and communicated, is proper only in the domain of mathematics and extension, but that the true estimation of real interactions requires that we shift to the Leibnizian, i.e. dynamic perspective, which speaks of internal forces, and is based on metaphysical concerns. He argues for this, however, by showing that mechanics and dynamics differ not just in their admission of intrinsic forces, but also in the fact that only the latter can understand the specific productive contribution of time to the development of force. The internal force of which Kant speaks is that which allows for an internal production of force that is not already contained in, and therefore reducible to, the antecedent motive causes of the system. In this manner, the consideration of evolution over time dissolves the strong link between the mathematical and the metaphysical, the quantitative and the qualitative, the external and the internal, the instantaneous and the temporal that held Leibnizian Dynamics together, but that equally became the noose around its neck. The young Kant, for all his misunderstandings, his misgivings, his misspeakings, at least attempted to dissolve this link, not to save the sanctity of the internal (towards which he will always take a

hesitant stand), but to save the importance of becoming (towards which he frequently turned his attention).

It seems that, in spite of all the refinements and the important new respect for Newton, Kant's project in the 1750s lies along the same lines: it seeks to make way for proper becoming in dynamics. This surfaces, for instance, in both of his major 1755 works, the *Universal Natural History* and the *Nova Dilucidatio*, and in his 1756 *Physical Mondology*. I will briefly comment on the contents of the latter two works in the remainder of this subsection, and look to the *Universal Natural History* in more detail in the following section.

In many ways, the *Principiorum primorum cognitionis metaphysicae nova dilucidatio* does exactly what we would expect of Kant on the basis of the *True Estimation*: an analysis of the precise implications of the principle of sufficient reason. Such an analysis is obviously meant as a criticism of the doctrines of German *Schulphilosophie*, which was based on this principle. In the tradition of philosophy inaugurated by Leibniz and Christian Wolff, substances do not really causally interact, and all the changes that occur in them are entirely due to the internal striving of their living force. According to such a view, all latter stages are already included in the essence of a thing, and only the temporal striving of the system also brings to the fore these pre-established features of the entity. The problem with this view is that it becomes unclear how there could be change, since the principle of sufficient reason requires that the change in internal state, however predetermined, be grounded in another change of conditions. For Leibnizians, these conditions would itself have to be internal, which means that the change of internal states is itself caused by an already effected change of internal states, which is an internal contradiction (AA I 410).

Kant seems to have been dissatisfied with a view that makes change and time a mere epiphenomenon of the independent, completely determined essences of things. Against it, he proposes a different view that change is both basic and causally relevant. For him, the internal changes to substances are due to the external changes in their situations, and not the other way around. Sure, substances can contain internal determinations that are then brought forth on occasion of a change in external circumstances, but they cannot themselves contain the ground of the change. Eric Watkins has raised a possible "charge of vacuity" to Kant's argument of succession in the *Nova Dilucidatio*: "it might seem that explaining the change of internal determinations on the basis of changing external grounds is vacuous insofar as one is still invoking change in order to explain change" (Watkins 2005: 132). I believe this charge can be resolved by interpreting Kant as stating that change and temporal evolution is fundamental to the universe, and not an epiphenomenon of the internal, determinations of things. On such a reading, Kant interpreted the Wolffians as stating that the principle of change is itself an internal determination of things, and as such itself a determination rather than a ground of determination. His criticism would then

be that change is a ground of determinations and of the possibility of mutual determination, rather than itself a determination/determined aspect. Prior to the actual course of the universe and the interactions occurring in it, the dynamics of individual entities cannot be determined.

There seems to be an important difference between this account and that of the *True Estimation* for our present purposes. The *True Estimation* argued that merely external determinations cannot account for the dynamic features of the real behaviour of bodies, and that there therefore need to be internal principles as well. In the *Nova Dilucidatio*, the stress goes the other way around: he argues that the internal principle of change does not of itself suffice to explain the dynamic features of the world. Of course, these two positions do not directly contradict each other. They may instead be taken as a shift of stress that reveals Kant's commitment to the need of both internal and external principles to have a truly dynamical universe. Neither suffices to ground change on its own, but together they allow for genuine change and mutual determination. Kant's discussions reveal as much, since in the *True Estimation* he argued that an internal force cannot be vivified unless it takes in external forces, and that the internal force which has accumulated external force can then produce novel results. The importance of this will become clear over the course of this dissertation.

The 1756 paper *Metaphysicae cum geometria iunctae usus in philosophia naturali, cuius specimen I. continet monadologiam physicam*, generally known as the *Monadologia Physica*, concentrates on the other aspect of the *True Estimation*, i.e. the metaphysical substrate of a genuine dynamics. One of the important goals of this essay is to reestablish the pan-elastic conception of nature (AA I: 486-487), this time on the basis of the two core powers of impenetrability (later repulsion) and gravity (attraction). Kant may have picked up on the fact that Boscovich's theory allowed for a reintroduction of Leibnizian elasticity in nature, thereby allowing for the beneficial aspect of the law of continuity that strengthened the causal links within nature. It also provides for a picture of the world that is more in line with that of the *Nova Dilucidatio*, namely of a universe of really interacting entities the intrinsic properties of which their responsiveness and resistance to external determination and change.

My concern here has not been to develop a full picture of Kant's early metaphysical physics, but rather to stress a central theme in it that is not often clearly discerned, namely that of change. But this theme relates problematically with some other concerns of Kant's. As we saw, Kant was interested in Leibnizian dynamics both for being dynamic, and therefore allowing for a contribution of time to physical processes, whereas he believed mechanics abstracted from this. On the other hand, he saw that the principle of sufficient reason has distinct benefits in guaranteeing the causal and intelligible closure of the world. Unfortunately, there is a deep tension between these two commitments, because the principle of sufficient reason is often so strong that it no longer allows the temporal evolution of a system to make a distinct contribution to the

system and the processes occurring in it, and instead predetermines all states in the initial conditions. In the Leibnizian framework, the principle of sufficient reason requires all change is fully determined by pre-existing determinations, and not itself a source of determinations, which makes it overly deterministic. In considerations on the laws of motion, this tension could not really come to the fore. It was rather in Kant's reflections on the genesis of the cosmos and of living systems that it would become pertinent. In the next section, I will discuss the tension as it appears in Kant's cosmology, while I will devote the whole of chapter two to the genesis of life.

## 1.3 From Cosmogony to Chaismogony

### 1.3.1 The Problem of Cosmogony in Early Modern Science

In spite of the major role played by astronomy in the formation of what we would now be tempted to call modern science, one important aspect of the structure of the universe and the heavenly (and terrestrial) bodies remained for ever problematic. Indeed, amidst all the debates on the structure of the universe, scientists and philosophers rarely touched upon the matter of the *genesis* of the universe. This is often ascribed to the limits imposed by the religious context in Early Modern Europe, namely Christian adherence to the narrative of the book *Genesis*. Any rival account of genesis faced the inevitable calumnious charges of heresy. Yet, this explanation, which is firmly built on the concept of the Science Wars, overlooks the important philosophical and metaphysical challenges this science faced; for on many accounts, genesis was not so much heretical as unintelligible.

This makes it all the more important to take a look at the major exceptions to this flawed rule. In fact, the earliest phases of modernity *did* offer alternative narratives of the generation of the universe, in the form of Descartes's *Le Monde* and a discussion in Pierre Gassendi's *Syntagma Philosophicum*. The work by Descartes is the better known and more influential, even though it appeared only after his death (in 1662 in a Latin translation, in 1664 in the original French), probably at least in part because Descartes briefly commented on it in the fifth part of his *Discours de la Méthode*. There, he mentioned the book in which he attempted to present his first reasonably comprehensive alternative physics, and said the following of its method of treating its great variety of controversial topics:

I did not want to bring these matters too much into the open, for I wished to be free to say what I thought about them without having either to follow or to refute



the accepted opinions of the learned. So I decided to leave our world wholly for them to argue about, and to speak solely of what would happen in a new world. I therefore supposed that God now created, somewhere in imaginary spaces, enough matter to compose such a world; that he variously and randomly agitated the different parts of this matter so as to form a chaos as confused as any the poets could invent; and that he then did nothing but lend his regular concurrence to nature, leaving it to act according to the laws he established. (AA VI: 42; Descartes 1985: 132)

Descartes here presents his account of the construction of the world in *Le Monde*, a work he would never publish himself, as a thought experiment in order to convince people of the explanatory potential of his theory. This presentation has often been suggested to be disingenuous, and merely an attempt to evade condemnation by the church. In a letter of from April 1634 (AT I: 284-289), Descartes intimated to his Paris correspondent that he was afraid to publish his *Le Monde* upon hearing of the condemnation of Galileo, with whom he agreed on the validity of Copernicanism. Given this general fear of persecution, Descartes is believed to have made an enormous amount of concessions in his work to religious authority, thereby frequently hiding his own views behind veils of implausible methodology.<sup>37</sup> This could lead us to believe that the presentation of the narrative of the genesis of the world as but a methodological or literary tool, is merely an attempt to mitigate the heretical connotations an alternative to *Genesis* carries with it. I do not find this plausible, however, and want to suggest that there are internal reasons why Descartes would not actually want to present his history of the universe as a real description.

First of all, if Descartes seriously believed his thesis to be heretical, he may have wanted to suppress it altogether from the *Discours de la Méthode* as well. On the other hand, if he believed presenting it as a mere methodological stance was enough to counter the charges of heresy, he could have believed his endorsement of Copernicanism could hide under the same shield. It is therefore likely that Descartes did not hold his claim regarding the Copernican structure of the universe and the Non-Biblical Genesis of the universe to be on a par.

If this is so, we may want to take Descartes at his word when he presents his own reasons for adopting this narrative in *Le Monde*. In particular, it allows him to introduce an important picture of the initial state of the universe, of the initial Chaos that reigned in it, namely that of an infinite continuous, purely extended matter devoid of the forms and qualities postulated by the scholastics, and hence containing “nothing which you do

---

<sup>37</sup> To name but two important philosophers and readers of Descartes who made a serious point out of this: Bernard Williams (2005) and Harry Frankfurt (2008) are both eager to identify the lasting effect of this fear of persecution in Descartes' beliefs or at least his presentation of them.

not know so perfectly that you could not even pretend to be ignorant of it.” (AA XI: 35; Descartes 1985: 91). In other words, the assumption of a new beginning allows Descartes to abstract from the complexities, the formations, the structures upon structures that many of his contemporaries considered to be the proper explananda of natural philosophy.

But besides this initial chaos, Descartes assumes two more aspects in the initial state: the totality of general laws of physics postulated by God and the articulation of the whole of extension in particular bodies through a certain division of motion and rest. Descartes then famously seeks to present a picture where the whole order of the visible universe arises through the behavior of mere extension as restricted only by the general laws of physics. Two important aspects of his narrative are how materials of different degrees of rarification are separated into distinct bodies of comparable density, and how the solar system is formed through the vortices constituted by the centrifugal forces acting upon the circling distinct bodies. Through these steps, Descartes hopes to explain primarily how planets are held together despite centrifugal force, and how the relative trajectories of heavenly bodies are possible only on assumption of mechanical effects.

There is however, here as elsewhere, a problem with Descartes’s genetical narrative, namely that it either asks too much of the initial conditions, or that it smuggles further determinations into them. The first problem is that Descartes seems to allow his general laws to do much more structuring work than seems warranted, for it is unclear why the world would not have evolved to a giant vortex where all the solid matter revolves around the centre and all the aether spirals around it. Moreover, he does not give a quantitative account to back up his rather vague fable. The second problem is that Descartes’s ordered universe is only possible on the basis of a specific set of initial conditions. Thus, God would have had to order the chaos in such a manner that specifically *this* universe and no other (with other planets, and for instance no inhabitable ones) resulted from it. We may want to conclude, therefore, that Descartes’s fable is more about the legitimacy of mechanical reasoning – which is analogical reasoning – than about the origin of the universe, and that its author admits as much by stressing that it is merely a fable. Indeed, for Descartes, all we need in order to explain reality is a specific geometrical structure and a specific set of general laws – but we cannot go about with only a random geometrical structure and a specific set of general laws, and hope that the latter will do all the structuring required.

The implausibility of the Cartesian narrative would soon reveal itself to the learned public upon the publication of the third part of Newton’s *Principia*. In the seminal general scholium to his monumental study, Newton attacked once more Descartes hypothesis of vortices, and in one and the same move ruled out the possibility of order arising out of chaos:

The hypothesis of vortices is beset with many difficulties. If, by a radius drawn to the sun, each and every planet is to describe areas proportional to the time, the periodic times of the parts of the vortex must be as the squares of the distances from the sun. If the periodic times of the planets are to be as the  $3/2$  powers of the distances from the sun, the periodic times of the parts of the vortex must be as the  $3/2$  powers of the distances. If the smaller vortices revolving about Saturn, Jupiter, and the other planets are to be preserved and are to float without agitation in the vortex of the sun, the periodic times of the parts of the solar vortex must be the same. The axial revolutions of the sun and planets, which would have to agree with the motions of their vortices, differ from all these proportions. The motions of comets are extremely regular, observe the same laws as the eccentric motions into all parts of the heavens, which cannot happen unless vortices are eliminated.

The only resistance which projectiles encounter in our air is from the air. With the air removed, as it is in Boyle's vacuum, resistance ceases, since a tenuous feather and solid gold fall with equal velocity in such a vacuum. And the case is the same for the celestial spaces, which are above the atmosphere of the earth. All bodies must move very freely in these spaces, and therefore planets and comets must revolve continually in orbits given in kind and in position, according to the laws set forth above. They will indeed persevere in their orbits by the laws of gravity, but they certainly could not originally have acquired the regular position of the orbits by these laws. (Newton 1999: 939-940)

Here, Newton argues that the general laws of nature alone could not, contrary to what Descartes seems to have argued, brought forth the current harmonic and equilibrated celestial system from just any initial configuration of matter. This suggests according to him that the configuration of matter (or at the very least the initial conditions of the current configuration of matter) need to have been selected specifically for the purpose of this harmony of the celestial system. Such a claim crippled the feasibility of a scientific treatment of the genesis of the universe, since it is only from highly specific initial conditions that the specific structure of the solar system can arise. In making astronomy a "proper" science at last, Newton had also relegated cosmology to the status of empty theorizing or mere storytelling.

### 1.3.2 Newtonian Cosmology? Nomology and Analogy

Despite Newton's authoritative dissuasion of rational cosmology, several thinkers of the High Enlightenment would venture into this hazardous discipline, undeterred by the warnings of one of the major *Lumières*. The reason for this is undoubtedly the context of the *Scholium Generale*, which starts off with a sustained piece of natural theology. From the constatation that the order of the universe could not have arisen otherwise than in

a single moment, Newton inferred that: “This most elegant system of the sun, planets and comets could not have arisen without the design and dominion of an intelligent and powerful being” (Newton 1999: 940). This pairing of the argument for constancy with that for intelligent creation would become an important force on the continent, for instance in Voltaire’s 1738 *Éléments de la Philosophie de Newton*:

The whole of Newton’s philosophy leads necessarily to the knowledge of a supreme being, who has created and arranged everything freely. For, if the world is finite, if there is a void, then matter does not exist necessarily, and has therefore received its existence from a free cause. If matter gravitates, as has been demonstrated, it does not seem to gravitate of itself, since it is extended of itself. Therefore, it has received gravity from God. If the planets revolve in one sense rather than another, in non-resistant space, the hand of the creator has thus directed their paths with absolute liberty. (Voltaire 1782 : 18; my translation)

Voltaire does not hesitate to indicate that Newton’s resistance to Descartes’s conception of the universe is invaluable in the battle against atheism, specifically Spinozism. Even though he believes Descartes himself to be innocent of this charge, he notes that

[t]he Cartesian system has produced that of Spinoza . [...] I have known many whom Cartesianism has led to admit no other God than the immensity of things, and I have on the contrary seen no Newtonian who was not theist in the most rigorous sense.

From the moment on that one is persuaded, with Descartes, that it is impossible that the world be finite, that motion is always preserved in the same quantity ; from the moment on that one dares say, give me movement and matter and I’ll give you a world; then these ideas seem to exclude, by valid inferences, the idea of a sole infinite being, a sole author of movement, a sole author of the organization of substances. (Voltaire 1782: 18-19; my translation)

It is the threat of Deism, of Materialism, of Mechanicism, of Spinozism, of Atheism, terms which turn out to be curiously equipollent in the mouths of many, that Voltaire seeks to combat by invoking Newton’s great scientific achievement. This achievement, Voltaire seems to say, implies that God is both the source of movement and the source of the organization, of the arrangement of matter in the universe. He contrasts this with the view he ascribes to Descartes, namely that God need only be the source of movement in the universe, and that the universal laws of the translation of movement

would suffice to bring about, from any initial arrangement of manner, the current harmonious structure of the world.<sup>38</sup>

Yet, Voltaire's idea that the belief in the productive force of the laws of nature leads automatically to Deism and Atheism, would draw the criticism of his rival in the divulgence of continental Newtonianism: Maupertuis. In spite of his championship of Newton in many things, Maupertuis does not hesitate to discount Newton's infamous proof in his 1750 *Essai de Cosmologie*:

The alternative of a choice or an extreme coincidence is founded only on Newton's impotence to give a physical cause for this uniformity. For other philosophers who let planets move in a fluid that drags them along, or that merely moderates their movement, the uniformity of their paths does not appear inexplicable at all. It no longer presupposes this singular coincidence or this choice, and no longer proves the existence of God any more than any other movement imprinted upon matter would. (Maupertuis 1768a: 9; my translation)

This is not to say that Maupertuis regards himself any less warranted to assume the existence of a higher being than the orthodox Newtonian. But instead of seeing the proof for his existence in the complex phenomena that result from the laws of nature, he sees it in the simplicity and the universality of the laws themselves (Maupertuis 1768a: 23). The difference here is between two conceptions of the hand of God. According to Voltaire, the existence and power of God is proved by the impotence of nature to produce the variety of harmonious effects occurring in it. According to Maupertuis, God's wisdom and power is instead proved by his ability to create a world that is itself productive and can produce complex and harmonious effects merely through general principles and laws, without having to interfere in the specific configuration of matter.

Maupertuis' more radical perspective on the productivity of nature, and therefore on the possibility of a scientific cosmogony, was shared by Georges-Louis Leclerc de Buffon. In the first discourses of his monumental *Histoire Naturelle*, Buffon presented both an iconoclastic methodology and an audacious theory on the evolution and formation of the solar system and the physical geography of the world. He ventured to formulate a theory of these controversial theories, however, only because he believed himself to be in the possession of an epistemology that allowed for qualifications. Indeed, instead of extolling the great synthesis of mathematics and physics that characterized the course of 17<sup>th</sup> century science, Buffon pointed towards the dangers of extending this method beyond its original domain, and separated mathematical and physical "truth":

---

<sup>38</sup> In section 4.1, I will elaborate on the deistic theory that God is the source of the arrangement or disposition of matter.

There are several species of truth, and it is customary to place the truths of mathematics in the highest order. However, these are but definitional truths which pertain to simple, though abstract, axioms, and all truths of this kind are but composite, yet always abstract, consequences of these definitions. [...] There is therefore nothing in this science other than that which we have placed in it ourselves, and the truths which we draw from it cannot be anything other than different expressions of the axioms that we have employed - so mathematical truths are but exact repetitions of its definitions or axioms. [...]

Physical truths, on the other hand, are never arbitrary and never depend on us. Instead of being grounded in the axioms we laid down, they rest only on facts. A sequence of similar facts, or, if you will, a frequent repetition and uninterrupted succession of the same events, is the essence of physical truth : that which we call a physical truth is therefore but a probability, but a probability so great that it equals a certainty. (Buffon 1749a: 54-55 ; my translation)

In these passages and those that follow them, Buffon sketches an anti-mathematical empiricism that believes mathematics to be certain but trivial and hence barren, and physics to be essentially association through resemblances and analogies that ultimately fails to yield actual causes, and hence gives us at best probabilities, however great these probabilities may be. Nonetheless, Buffon believes the two can be fruitfully applied to each other, albeit only in very limited cases. His sole example of such a limiting case should not come as a surprise:

The most beautiful and fortunate such application that anyone has ever made, is to the system of the world; and it has to be admitted that, if Newton had given us only the physical ideas of his system, without having supported them with precise evaluations and mathematics, they would not have had nearly the same force. But at the same time we must realize that there are very few subjects so simple, that is to say, so bereft of physical qualities, than that one, since the distance between planets is so great that we can consider them with regard to each other as simply being points. We can also, at the same time, without erring, abstract from all physical qualities of the planets, and consider only their force of attraction. Furthermore, their movements are the most regular that we know, and exhibit no slowing due to resistance. All this taken together makes the explanation of the system of the world a mathematical problem, which requires only a fortunate physical idea to be realized. This idea is that the force that makes thing fall towards the surface of the earth, could very well be the same as that which keeps the moon in its orbit. (Buffon 1749a : 58-60)

Buffon seems convinced that for most inquiries into nature, the approach typical of mathematical physics will prove unfruitful or unpracticable or both. He explains the success of mathematical physics not by the universal applicability and desirability of its method, but rather by the physical and theoretical simplicity of the systems it

considers. Here we see a first version of the idea that classical mechanics is tailored to an overly simple universe, and fails from the moment we seek to apply it to more complicated problems. Buffon makes these claims in the preliminary discourse to a work that is mainly on natural history, that is to say, the study of what we would now call biological systems and their closest analogues in nature. He thereby submits that, in the case of these latter systems, mathematical physics will only lead us astray, and cause us to overlook the complexity of the phenomenon in our quest for abstraction. For him, a proper method must seek to grasp this complexity, and reason through actual effects and causes in order to acquire a theory, the probability of which can then be estimated mathematically.

But the first discourses of the *Histoire Naturelle* are not on living systems, but on the genesis of the celestial system and of the geological and geographical composition of the world. Buffon justifies his theorizing in those tracts by noting that, in the former problem, we are dealing with a very limited amount of basic forces exerted on simple mechanical systems (Buffon 1749a: 131), and in the latter, we have an abundance of evidence at our disposal (Buffon 1749a: 68). Even then, however, complications inevitably arise, such as the complexity of the solar system, and the origin of the original velocity of the planets used to explain their velocity. In spite of, or perhaps because of their boldness, the major cosmogonies of the mid-eighteenth century remained vague and sketchy in their conjectures. They failed to show how exactly such general forces would give rise to a harmonic world system, thereby seeming to beg the question against those who remained skeptical until the vagueness was completely dispelled.

### 1.3.3 Time and Teleology

It is far from a coincidence that, in the 1750s, Kant wrote a work of conjectural cosmogony titled *Universal natural history and theory of the heavens or essay on the constitution and the mechanical origin of the whole universe according to Newtonian principles*. In this essay, published in 1755, he attempted, like Buffon, to present a scientific explanation of the structure and origin of the solar system, and, like Buffon, he remarks that it is because of the generality of the laws of physics and the simplicity of spherical bodies like planets as a guarantee that such an attempt has a chance at succeeding (AA I: 229). The parallels with Buffon are far from coincidental, since the work of the French natural historian has been noted as a major influence on Kant throughout his life, especially in the 1750s (Mensch 2013: 59).

Like Buffon and Maupertuis, moreover, the overt Newtonianism is balanced by a more Cartesian conception of a decent cosmology. As William Shay (1986: 105) noted, “Kant's cosmology is more in the spirit of Descartes' *Principia Philosophica* [sic] than

Newton's *Principia Mathematica*. The chain of reasoning is not held together by rigorous mathematical links but by appeals to analogies". This should not be a surprise, since we already saw in 1.1.2. that Kant is more properly understood within the tradition of natural philosophy in which Descartes and Leibniz figured than as placed in the more straightforwardly scientific tradition in which we can place Newton. Thus, Kant is not so much in the business of providing a mathematical cosmology rivalling that of Newton, as of showing, like Maupertuis and Buffon, that on Newtonian principles, the genesis of the universe according to general laws, could occur, contrary to Newton's own insistence.

There are several reasons why Kant believes, or could have believed, himself to be entitled to analogical reasoning in such an account. First, like Buffon, Kant seems to have believed that mathematical infallibility can never be demanded of a treatise of this kind. If the system is based on analogies and harmonies in accordance with the rules of credibility and a correct way of thinking, it has satisfied all the requirements of its object. (AA I: 235) Second, the reasoning from analogy seems justified, as Shay (1986: 106) has stressed, by the principle of continuity, which allows us to regard otherwise distinct phenomena as sharing important properties.<sup>39</sup> Finally, Kant's analogical argument can suffice for the present purpose because, as Eric Schliesser (2013: 426) has indicated, he needn't show *how* Newtonian principles give rise to the current world, but only *that they can* do so. Like Descartes, and like Buffon and Maupertuis, all Kant requires is a "how plausibly" account, not a real account, to weaken Newton's radical conclusion of divine design.

If we want to assess Kant's approach here, we must turn to his account in the work itself, the explanandum of which is the "systematic constitution of the universe":

Actually, all the planets and comets that belong to our universe constitute a system simply because they orbit around a common central body. But I take this term in a narrower meaning in that I consider the more precise relationships that have made their connection to one another regular and uniform. The orbits of the planets relate as closely as possible to a common plane, namely to the extended equatorial plane of the Sun; the deviation from this rule occurs only at the outermost border of the system, where all motions gradually cease. If, therefore, a certain number of heavenly bodies that are arranged around a common central point and move around this, are simultaneously restricted to a certain plane in such a way that they have the freedom to deviate from it to either side only as little as possible; if such a deviation occurs gradually only in those that are most remote from the centre point and thus participate less in the relationships than

---

<sup>39</sup> In 5.4.1, we will see this connection of the principle of continuity with that of analogy, and the necessity of analogy, resurface in Kant's discussion of the principle of reflective judgment.



the others: then, I say, that these bodies are related to each other in a *systematic constitution*. (AA I: 246)

In the first part of the *Universal Natural History*, Kant attempts to show, by reasoning from analogy with the available astronomical data, that “[t]he fixed stars, as we know, all relate to a common plane and thus constitute an orderly whole, which is a world of worlds. One can see that in the immeasurable distances, there are more such star systems, and that creation in the entire infinite scope of its size is everywhere systematic and interrelated”, and speculates “that these higher orders of worlds are not without connection to one another and that, through this mutual relationship, they constitute in turn an even more immeasurable system” (AA I: 255). But besides reasoning through analogy, there is another principle which allows him to entertain this hypothesis: “[i]f the parts of nature are observed according to intentions and a discovered plan, certain properties are revealed that would otherwise be overlooked and remain hidden if our observation is spread over all other objects without any guidance” (AA I: 255). This additional guiding principle is teleological, in that it states that we must look to the universe with an eye for systematicity in order to appreciate its constitution. Besides systematicity, however, Kant points towards infinity as a feature brought to light by a teleological perspective on nature:

The theory we have put forward opens a perspective onto the infinite field of creation for us and presents some inkling of God’s work that is appropriate to the infinitude of the great architect. [...] We see the first members of a progressive relationship of worlds and systems, and the first part of this infinite progression already gives us to understand what we can suppose about the whole. There is no end here but rather an abyss of a true immeasurability into which all capacity of human concepts sinks even if it is raised with the help of mathematics. The wisdom, goodness, the power that has revealed itself, is infinite and in the same measure fruitful and industrious; the plan of its revelation must for that reason be as infinite and without limits as it is. (AA I 255-256)

The universe is characterized, besides by its systematicity, by an immeasurability that challenges our conceptual grasp. In chapter 7, I will show that this tension between the systematicity and the immeasurableness of nature are fundamental to the doctrine of the sublime in the *Critique of the Power of Judgment*. But now we must first see how Kant believes such an orderly yet immeasurable universe to be able to come into existence at all.

Like Descartes before him, Kant begins by assuming that the universe was chaotic at its inception, and that the materials out of which the heavenly bodies exist were distributed evenly over the space now occupied by the systems in which they are organized. But just as Kant has described this initial state, where the universe is “as raw, as unformed as possible” (AA I: 263), he goes on to add a non-trivial caveat:

However, even in the essential properties of the elements that make up chaos, the characteristic of that perfection can be felt that they have from their origin, in that their essence is a consequence of the eternal idea of divine reason. The simplest, the most universal properties that appear to have been designed without any intention, matter that seems to be merely passive and in need of forms and arrangement, has, in its simplest state, an endeavour to form itself into a more perfect state by a natural development. However, the *difference in the kinds of elements* contributes the greatest part to the regulation of nature and the formation from chaos by which the state of rest that would prevail under a universal equality among the dispersed elements, is eliminated and the chaos in the points of the more strongly attracting particles begin to form. (AA I: 263-264)

In this passage, Kant admits that two additional elements are needed to get order out of chaos, i.e. to have a universe evolve from a different configuration to its current harmonious and systematic one. The first of these elements is that matter cannot be merely passive, but must instead be admitted to have an internal endeavor to form and articulate itself even in its most simple state. This first suggestion echoes the *True Estimation*, where Kant suggested that we must postulate some internal force in the body undergoing change in order to understand how, over time, it comes to acquire properties that were not fully in its initial conditions. Initially, we might take Kant's talk of essential forces and sources of life as merely metaphorical, given that he may very well be talking about the force of attraction, since he passes to a description of the formation of lumps of matter around the more dense elements. But these lumps "would, after they had completed their formation, remain at rest and eternally unmoving because of the equality of attraction" (AA I: 264). Kant resolves this problem by attributing, besides the attractive force, a *repulsive* force to nature, which "by [its] conflict with the attractive force, bring[s] about that motion that is, as it were, a continuous life in nature" (AA I: 265). He suggests, then, that the universe can be understood to be dynamic and genetically productive if we understand matter as subject to two opposing internal forces, and not just as a ragdoll being tossed about by external, mechanical forces.

But Kant adds that another thing is needed for the emergence of order out of chaos, and thereby reveals that his initial state of chaos of passive matter is not quite as chaotic as it appeared to be. In fact, Kant denies that the universe could be, in its initial state, homogeneous, since this would make it impossible for any motion to occur. If matter were homogeneous and evenly distributed in an infinite space, the mutual attraction between all elements would keep the whole universe in a static equilibrium. The initial state of Kant's universe, however, is one of heterogeneous matter, whereby a variety of kinds of elements each have different intensional properties: "[i]n a space filled in such a way, universal rest lasts only a moment. The elements have essential

forces to put each other into motion and they are a source of life for themselves” (AA I: 264).

In his stipulation of the initial state of the universe, Kant shows himself very sensitive to the challenges to any physical cosmogony, understood as a theory of the genesis of a universe that is orderly rather than chaotic. One of these challenges is that a homogeneous initial state will never of itself become dynamic: it will remain forever in a state of static equilibrium. The other is that the principle of change invoked should not have as effect an evolution towards a state of static equilibrium. In order to get systematicity and dynamics into our universe, we have to stipulate it as somehow already present in the initial state. On second thought, then, Kant’s initial chaos is not all that chaotic: it is already replete with the orderliness and productive force that it should itself give rise to.

There is also another way in which the universe presented here is not homogeneous: it has a center, in casu a center of gravity. As Kant himself writes:

It is certainly true that in an infinite space, no point can properly have the prerogative of being called the centre point; but by means of a certain relationship that is based on the essential degrees of the density of the original material, according to which, at its creation, this is initially more densely concentrated at a particular place and increases in its dispersion with distance from that place, such a point can have the prerogative of being called the centre point and it actually becomes through the formation of the central mass of the strongest attraction therein, to which all the remaining elementary matter that is in the process of coalescing into particular formations descends and thereby, however far the evolution of nature might extend, makes just a single system out of the whole of the universe in the infinite sphere of creation. (AA I: 312)

As this picture progresses, the Kantian universe slowly loses its chaotic, homogeneous, random nature, and becomes one which was structured, in advance, to yield the kind of orderly universe in which we live. It is topologically heterogeneous, since it has a common centre of gravity. It is chemically heterogeneous, because it consists of a wide variety of elements with different properties (most importantly, different specific masses). Finally, it is characterized by a duality of general principles that counteract each other’s homogenizing and equilibrating effects. With these stipulations, Kant shows himself eminently aware of the fact that physics itself leads us to marvel over the specificity of the universe, such that it could have led to the dynamic orderly pattern that it is today. In this realization, he agrees with Albert Einstein, for whom the cosmological problem was so vexing.

This raises the following question: does Kant’s account of the genesis of the universe still invoke teleology, or is it another species of the atheism, spinozism and epicureanism that Voltaire identified as the logical outcome of the idea of mechanical genesis? That Kant was aware of this difficulty is evinced by the fact that Kant repeats,

at two occasions in the preface, Voltaire's phrase "give me motion and matter, and I'll make you a world", and that he presented it as a possible charge of epicureanism on each of those occasions. Against this charge, he advances two arguments that are to be expected in the light of Maupertuis' *Essai de Cosmologie*. The first is that he believes himself to be expressing great esteem for the omnipotence of God in stating that with mere matter and motion he can allow order to arise out of initial chaos (AA I: 228). The idea is that Newton and Voltaire commit *lèse-majesté* when they suggest that God could not have created a genuinely productive nature. The other is that he merely adopts a different picture of teleology. Recently, Martin Schönfeld (2000: 98) has argued against the wide-spread view that the *Universal Natural History* is an anti-teleological work. His basic claim is that Kant's teleology is of a radically different nature from that of many of his predecessors and contemporaries, since it is anti-anthropocentric, but that it is nonetheless a form of teleology, and Huneman (2008: 93) has equally noted that Kant's early teleology is of a different kind than Newton's. This argument, however, does not yet figure importantly in the *Universal Natural History*, and becomes apparent only in the passage of a 1763 paper where he revisits the teleological implications of his own earlier work:

If we discover an arrangement in nature, which seems to have been instituted for a special purpose, since the general properties of matter on their own could not have produced such an order, then we regard this provision as contingent and as the product of choice. Now, if new harmony, order and usefulness should make their appearance, along with mediating causes especially instituted to produce these effects, then we judge them in the same way to be contingent and the product of choice. This connection is quite alien to the nature of the things of themselves. They stand in this harmonious relation simply because someone has chosen to connect them in this way. No general cause can be adduced to explain the sheathed character, that is to say, the retractability of the claws of the cat, the lion and the so on. The only explanation which can be given is that a Creator has ordered them in this way, with a view to protecting them from wear, for these animals must have implements suitable for seizing and retaining their prey. But suppose that matter has certain properties of a more general character, which in addition to producing certain benefits which may be construed as their *raison d'être*, are also particularly suited to producing even more harmony, and doing so without the least provision being made to bring it about. Suppose that a simple law, which is universally agreed to be necessary for the production of a certain good, also produces fruitful effects in many other ways as well. Suppose that a simple law was the source of further usefulness and harmoniousness, not by art, but rather of necessity. And suppose, finally, that this should hold throughout the whole of material nature. If all this were supposed, then there would obviously inhere in the very essence of things themselves universal relations to unity and cohesiveness, and a universal harmony would extend throughout the realm of

possibility itself. Such a state of affairs would fill us with admiration for such extensive adaptedness and natural harmony. Adaptedness and natural harmony such as this, although rendering punctilious and forced art superfluous, can nonetheless never themselves be ascribed to chance. It rather indicates that there is a unity to be found in the possibilities of things themselves: it suggests that the essences of all things are without exception dependent upon one single ground. (AA II: 96-97)

Kant suggests here that the real opposition is not between teleology and chance or necessity, but rather between the anthropocentric or anthropomorphic teleology that expects all harmonious arrangement to be directly intentionally designed, for instance by the hand of God, and a natural teleology which regards nature as such a unity and productive structure that it can give rise to harmony. In chapter 4, we will see that this contrast between intentional and non-intentional teleology emerges as well in the *Critique of the Power of Judgment*, and in chapter 5 we will see why Kant still insists on calling “purposive” the unitary structure of nature and its various laws.

But the passage just quoted differs from Kant’s position in the *Universal Natural History* in suggesting that the structure of organisms too might be the product of such a productive nature. In the *Universal Natural History*, Kant still denied that more complex entities, organisms in particular, can be explained on the mechanical grounds posited here (AA I: 230). In the next chapter, we will see how Kant’s views on the genesis and generation of animals evolved such that he could regard them as genuinely naturally generated and produced rather than intentionally designed by God.

## 1.4 Conclusion

In this chapter, I have argued that in Kant’s early engagement with the natural sciences, he was already concerned with the limitations of Newtonian mechanics and Leibnizian dynamics rather than with celebrating the victory of either of these evolutions in the natural sciences. In this conclusion, I will try to spell out more explicitly what these limitations are. Both the *True Estimation* and the *Universal Natural History* seem concerned with the following problem: how can we attribute a genuine dynamic nature to the universe given the constraints of physical interaction. The problem can be elaborated as follows: the rigour of both Newtonian mechanics and Leibnizian dynamics seems to lie in the strict connection they postulate between the initial state and the subsequent state. I have tried to show that Kant was dissatisfied with Leibnizian dynamics because it did not allow for actual change in the universe, and pushed instead for an account that could allow for genuine change to occur in a universe that is physically intelligible. I

have also argued that a similar problem arose in the *Universal Natural History*, where Kant sought to demonstrate that a Newtonian system need not be made of one piece, but can form itself over time from another system. It is important, however, that Kant remained aware of the rationale behind both Leibniz' and Newton's ideas. This rationale was that in fact, there are substantial constraints on the creative power of physical nature. It remains impossible to get order to arise from chaos. If it seems as if Kant is denying this latter statement, it is because his chaos is not ours: it is a very structured chaos indeed. In the *Universal Natural History*, it becomes clear that the problem of the constraints on the productive force of nature relates to the problem of teleology. After all, if the universe can give rise to order, to complexity, only if it already contains a great deal of order complexity, then its productive capacities are already highly determined at its inception. Secondly, Kant argues, in both works, that we must ascribe more powers to nature if we want to understand how it can be genuinely productive. In the *True Estimation*, he postulates an internal force that is capable of internalizing externally impressed forces in order to produce effects that were not already fully included in the causes. In the *Universal Natural History*, he stresses that matter cannot be merely passive, and must be endowed with internal striving forces in order to be capable of generating harmony out of chaos. These reflections must remain vague until they are applied to a domain where they can become clear-cut. This domain is that of the genesis of living beings, where the problem is pushed to its extremes in the idea of equifinality. An equifinal process is a process that leads to the same outcome in spite of (a specific range of) differences to the initial state. It should not come as a surprise, therefore, that in subsequent phases of Kant's thought, the problem of animal generation became more and more central.

## Chapter 2

# Epigenesis and the Dynamic Nature of Life: Kant's Theory of Generation

*L'emblème antique qui représente la vie par un cercle formé par un serpent qui se mord la queue donne une image assez juste des choses. En effet, dans les organismes complexes, l'organisme de la vie forme bien un cercle fermé, mais un cercle qui a une tête et une queue*

- Claude Bernard

In the previous chapter, I have argued that even in Kant's engagement with physics, the problem of teleology was of central importance, mostly insofar as it is connected with the problem of the genesis and development of structured systems. This problem for early modern science has been usually discussed not in connection with physics, but with the life sciences, more specifically the theory of generation. In this chapter, I will argue that Kant developed, over the course of several decades, his own version of the theory of epigenesis in response to the problem of the genesis of structured systems. I will be arguing against the now wide-spread assessment that Kant was ultimately committed more to a mitigated form of preformationism rather than epigenesis, and that he therefore denied the epigenesist stress on historical development as truly productive of novelty rather than merely a series of occasions for the expression, of preestablished, innate principles. In section 2.1, I present the view against which I will be arguing. My argument takes two phases. In section 2.2, I argue against the idea that Kant faced a dilemma between epigenesis and preformation, and for the idea that he was faced with three options rather than two. I present these three options in subsections 2.2.1, 2.2.2 and 2.2.3. In subsection 2.2.4, I present the revival of epigenesis in Kant's own time. In section 2.3, I argue that Kant is already best understood as an epigenesist in his pre-critical period, both in his 1763 book (2.3.1) and his 1770s essays (2.3.2), and that his preformationist language there is subservient to the idea of epigenesis as that of the natural productivity of life. In section 2.4, I argue that this

position remained stable throughout the critical period. In 2.4.1, I argue that we cannot take Kant's rejection of Herder's version of epigenesis to be a refusal of epigenesis altogether, because Kant believed that some internal constraints must be placed on the genetic force. This position also allows us to explain, in 2.4.2, Kant's endorsement of epigenesis in the *Critique of the Power of Judgment* and its description there as "generic preformationism". I conclude by listing the main features of the Kantian theory of epigenesis.

## 2.1 Kant and Early Modern Theories of Generation

The early modern debate between proponents of the rival theories of generation commonly known as preformationism and epigenesis has usually been viewed as a species of the larger recurrent debate between mechanist and vitalist approaches to life. This assumption appears, amongst others, in the image of Canguilhem's pendulum:

biological theory reveals itself to be a thinking that throughout its history has been divided and oscillating. Mechanism and Vitalism confront one another on the problem of structures and functions; Discontinuity and Continuity on the problem of the succession of forms; Preformation and Epigenesis on the problem of the development of a being; Atomicity and Totality on the problem of individuality. (Canguilhem, 2008, p. 61)

It is unclear from this passage whether Canguilhem himself believes these different dichotomies to map neatly onto each other.<sup>1</sup> But the enumeration and the parallelism suggests that he at least realizes many have thought this to be so. In the more recent literature, the assimilation of the preformationism-epigenesis debate to the mechanism-vitalism debate has only become more apparent. Shirley Roe (1981) and Jane

---

<sup>1</sup> The main reason for my doubt here is the position of the dichotomy between discontinuity and continuity, since it is far from obvious that continuity is always on the side totality and vitality. In fact, throughout history, continuity and discontinuity have been associated, in different periods and even in different theories of the same period, with opposite implications and characteristics. In early modernity, continuity suggested panorganicism to some, whereas it was proof for materialism to others. I suspect, however, that Canguilhem might be thinking primarily of the debate between mutationism and Darwinism around the turn of the 19<sup>th</sup> and 20<sup>th</sup> centuries, where mutationism was regarded as defending sudden formal change whereas Darwinism stressed the primacy of functional restrictions. But the choice between structure and function as the major motor of speciation does not necessarily prejudice one towards either continuity or discontinuity.



Maienschein (2000) have offered such an interpretation, the outline of which is the following:

The structure of a living system is deemed, by most early modern thinkers, too complex to be formable through mechanical means alone - “through mechanical means alone” meaning, of course, through mechanical processes involving essentially inert matter. Resolving the problem of generation subsequently prompts a choice between two positions:

- 1) **Epigenesis**, which maintains that the complex structure is formed gradually from more homogeneous matter, but that this process of formation is guided by a certain principle that may be either immaterial (as in vitalism as classically construed) or a property of the matter employed in generation (which is therefore no longer taken to be inert - this position amounts to hylozoism or vital materialism).
- 2) **Preformationism**, which holds fast to mechanical principles, and denies instead that the structure of the living system is genuinely formed during generation. It is important to realize that Preformationism does not necessarily require the preexisting structure to be a miniature of the generated structure. It does, however, require the preexisting structure to be as complex as the generated structure and to be able to yield the latter through mechanical transformations alone.

The early modern theorist of life must therefore either deny genuine formation and accept preformation, or abandon mechanist materialism and accept epigenesis.

This picture is distorted because it takes both positions to agree essentially on the fact that the generated structure must in some sense be pre-given, and to disagree only on whether and in what sense it must be pre-given *materially*. Its advocates may be motivated by the much later discovery of a material element - an acid, to be precise - believed to be endowed with a guiding structural property - the “information” of the “sequence” - to lie at the basis of the process of generation. This discovery is believed to resolve the ancient dilemma by doing justice to the core intuitions of both positions: the fact of genuine generation and the material preexistence of structure. Jason Scott Robert (2004: 37-38) has called the view that the discovery of DNA yielded a synthesis of sorts between epigenesis and preformationism the “modern consensus”. It can be discerned in authors as diverse and divergent as Jacques Monod (1970: 117), Ernst Mayr (Mayr 1977: 157-158) and Stephen J. Gould (1977: 8-19).

The idea that epigenesis and preformation ultimately converge is also nourished by the fact that some have taken them to address distinct concerns, with epigenesis

seeking to offer a description of the perceivable formation of a living structure, and preformationism seeking to explain this formation. Detlefsen (2006), for instance, claims that in the paradigmatic case of the Haller-Wolff-debate, Wolff's epigenesis moves considerably towards a preformationist position the moment it seeks to accommodate the desire for an *explanation* rather than a mere description of formation.

It is not clear, however, that the modern synthesis is a genuine compromise between the opposing positions. First of all, many authors are only tempted to regard it as non-preformationist because they identify preformationism with its naivest versions. As noted above, it need not be the case that the preformed structure is a *miniature* of the end-product, since most preformationists allow for transformations to occur during the process of development. Secondly, one can be tempted to regard the information of the DNA as the guiding principle, but this is notoriously problematic, since it tends to overstretch the proper meaning of the physical and mathematical notion of information, and give it quasi-intentional connotations.

I believe, along with for instance Susan Oyama (2000: 28-35), that the modern consensus is merely a mitigated preformationism. This explains, for instance, the fact that preformationism has been rehabilitated in recent years, and even regarded as an adumbration<sup>2</sup> of our current best theories, for instance by Clara Pinto-Correia, who writes:

The belief in the preprogrammed encasement of successive generations was the centerpiece of preformation. However, the organization of that encasement underwent several revisions. In the course of a century (from the mid-1600s to the mid-1700s), preformationist theories evolved from totally preformed persons to pre-existing fundamental parts. *In its final forms, the theory came one step shy of our current models in developmental biology.* But that step, in the context of the time, was impossible to take. (Pinto-Correia 1997: 8)

That this view has, in turn, determined our reading of Kant, is clear from the following passage from the end of Pinto-Correia's book:

Newly found systems such as the *wnt* and *TRK* pathways certainly show the 'truth' of epigenetic interactions on a more than descriptive level. But no one doubts that these molecules are genetically determined and placed in the appropriate cells through interactions of enhancers and transcription factors encoded by the

---

<sup>2</sup> The pun here is intended: throughout her book, Pinto-Correia uses a preformationist theory of the history of ideas to analyze the history of preformationism. In her own words: "human ideas are always the children of pre-existing ideas [...] whenever a new thought surfaces in someone's mind, it has to be a metamorphic form; a result of the continual rearranging and reconfiguring always going on deep within our mental subsoil, upon layer upon layer of sediment." (Pinto-Correia 1997: 5)

'preformed' genome. Kant and Blumenbach had already foreseen something like this, but we are still unpacking Wilson's 1925 notion that ontogenesis is a cytoplasmatic epigenesis underlain by a nuclear preformationism. (Pinto-Correia 1997: 310)

Pinto-Correia is not alone in this assessment of Kant as endorsing a middle way between preformation and epigenesis. In the early 2000s, Philip Sloan (2002) and John Zammito (2003) argued that Kant's usage of the term epigenesis to describe his own theory of generation should not be taken at face value. Instead, they believe Kant evolved from an unequivocally preformationist position, which he held in the 1760s, 1770s and early 1780s, to a position which was a middle way between epigenesis and preformation from the mid-1780s onwards. This reading is consequential because it has serious implications for the interpretation of Kant's use of terminology from the theory of generation in his philosophy of history and his epistemology, as we will see in chapter 3.

In this chapter, I will argue against Sloan and Zammito that Kant is best understood as endorsing epigenesis from the early 1760s onwards, and to view all later texts and elaborations on the topic as a further development of the position he envisaged early on in his career. I will do so by arguing that Kant is badly understood as having to choose between epigenesis and preformation, because, as I will show in section 2.2., to Kant and his contemporaries there were three major alternative groups of theories available to them: epigenesis, preformation and metamorphosis. This will allow me to identify the proper targets of Kant's criticism and praise in matters of embryology. I will also show that Kant's allegedly unambiguous endorsement of preformation in his early career is nothing of the kind, by developing, in section 2.3., a reading of these early works according to which they are thoroughly epigenesist. Finally, I will show, in section 2.4. that Kant's critical version of epigenesis is precisely that: a critical version of epigenesis. I will argue there that what Sloan and Zammito have taken to be an ambiguous endorsement of epigenesis is in fact meant to be a sensible version of epigenesis, the only reticence being that required by the epistemological stance that is transcendental idealism.

## **2.2 Structure and Genesis in Early Modern Theories of Generation**

In this section, I will sketch an alternative picture of the Early Modern debate on animal generation. In the first place, I will argue that is inadequate to present the debate as between two positions, namely epigenesis and preformationism. This idea of a

dichotomy between the two positions has led many authors to identify a variety of alternative forms of either theory. It is, for instance, common to argue that there is a distinction between mechanist and more vitalist or vital materialist versions of epigenesis. On the other hand, some authors distinguish between two versions of preformation, one according to which the living system has been structured from all times, and one according to which it is merely rapidly formed directly before or directly after conception. The peculiarity is that, according to such a schema, there are theories of generation that are epigenesist in some respects and preformationist in others. I will argue that this problem can be resolved by recognizing that these so-called intermediate positions are actually an autonomous third position, which I call, after William Harvey, metamorphosis.

Second, I will present this triad of preformation-metamorphosis-epigenesis as a trichotomy because each position entails a different answer to the problem of animal generation. This problem is that of explaining, or making intelligible, how the highly complex and (in the language of the time) contrived structure of living systems can come into being, given that any such account seems to involve a regress to a structure explaining such a structure. We will see that epigenesists respond to this regress by proposing that generation involves circularity, that preformationism takes recourse to a linear infinite regress, and that metamorphosists attempt to find a level at which the explanation can reasonably bottom out. In the course of explaining this, I hope to also explain why these positions can sometimes be mistaken for each other.

Third, I will show that each of these positions are refined and theoretically powerful. For all their shortcomings, they cannot be identified with the caricatures that we find in some textbooks and outdated overviews of the history of science. This is important because it will provide better access to Kant's own position. First of all, it will become clear that some of his comments that have been taken as criticisms of epigenesis are actually directed at metamorphosis. Second, it will become hard to argue, as Sloan and Zammito have, that Kant's criticisms were directed at an earlier, unrefined version of preformation, and meant to be in favor of a later, more refined version, once we realise that the unrefined version did not exist in the eyes of the 17<sup>th</sup> and 18<sup>th</sup> century. What we see as the unrefined version is a caricature developed later in history, in the narratives of developmental biology from the late 19<sup>th</sup> and early 20<sup>th</sup> centuries. Third, we will be able to see that Kant's characterization of epigenesis as generic preformationism is not a concession to preformationism, but instead precisely what it seems to be at face value, namely an insight into what kind of structural constraints are required by all epigenesist theories. With this comment, Kant is indeed distancing himself from the epigenesis of Herder, but allying himself with that of Aristotle, Harvey, Wolff and perhaps Blumenbach.

### 2.2.1 Epigenesis and the Circular Paradigm

Although the theory of epigenesis goes back to Aristotle, who proposed, in the *Generation of Animals*, the idea that living beings are formed (at least in the case of sexual production) from relatively homogeneous matter (e.g. menstrual blood) by an active principle (contained in the male semen), its seminal Early Modern version was formulated by William Harvey in his 1651 *Exercitationes de Generatione Animalium*. In that work, the man usually credited with the discovery of pulmonary circulation in the West set out to attack those anatomists who theorize "as if *generation* were nothing more than a separation, or aggregation, or disposition of things" (Harvey 1847: 207). This position, which he ascribes to most authors who worked in a broadly Hippocratic tradition, he named *metamorphosis*, and contrasted with his own proposal, or rather his own reappraisal of Aristotle's proposal, namely *epigenesis*. It is instrumental to my argument that I briefly linger on the precise distinction between these two positions, beginning with Harvey's interesting characterization of metamorphosis.

In the 45<sup>th</sup> of his *Exercitationes*, Harvey introduces metamorphosis as follows:

Some [animals], out of a material previously concocted, and that has already attained its bulk, receive their forms and transfigurations ; and all their parts are fashioned simultaneously, each with its distinctive characteristic, by the process called metamorphosis, and in this way a perfect animal is at once born. (Harvey 1847: 334)

The distinctive mark of metamorphosis is then the fact that it allows for the generation of animals, not as a gradual process, but rather as a instantaneous one, whereby the organism is instantaneously, or precipitately, formed out of its bare matter. In an earlier *exercitatio*, Harvey claimed that the proponents of such a theory assign the causes for this process mainly to the underlying matter: "they who philosophize in this way, assign a material cause [for generation], and deduce the causes of natural things either from the elements concurring spontaneously or accidentally, or from atoms variously arranged" (Harvey 1847: 207). Although he finds such an approach highly problematic for higher animals, he does not rule out that in this way,

The generation of insects is effected where by metamorphosis a worm is born from an egg; or out of a putrescent material, the drying of a moist substance or the moistening of a dry one, rudiments are created, from which, as from a caterpillar grown to its full size, or from an aurelia, springs a butterfly or fly already of a proper size, which never attains to any larger growth after it is first born; this is called metamorphosis[, wherein] chance or hazard seems the principal promoter of generation, and there, the form is due to the potency of a preexisting material ; and the first cause of generation is 'matter', rather than 'an external efficient'. (Harvey 1847: 334-335)

What we gather from this passage is that Harvey regards the theory of metamorphosis, which focuses on the material cause of generation, as an illegitimate extension of the process of spontaneous generation (not yet an inadmissible notion at that time) of relatively simple animals to higher orders of natural beings.

Against the defenders of metamorphosis, Harvey maintained that

There are some [animals] in which one part is made before another, and then from the same material, afterwards receive at once at once nutrition, bulk and form: that is to say, they have some parts made before, some after others, and these are at the same time increased in size and altered in form. The structure of these animals commences from some one part as its nucleus and origin, by the instrumentality of which the rest of the limbs are joined on, and this we say takes place by the method of epigenesis, namely, by degrees, part after part, and this is, in preference to the other mode, generation properly so called. (Harvey 1847: 334)

In contradistinction to metamorphosis, then, the process of epigenesis is gradual and ordered: it creates the different parts of the organism one after the other, starting with a specific part as origin, which then allows for the creation of further parts, which in turn allow for the generation of the specificity of yet more parts, until a fully articulated organism is formed. Also, it is not driven by the specific properties of the underlying matter:

An animal which is created by epigenesis attracts, prepares, elaborates, and makes use of the material, all at the same time: the process of formation and growth are simultaneous. [W]hile it creates in succession parts which are differently and variously distributed, it requires and makes a material which is also various in its nature, and variously distributed, and such as is now adapted to the formation of one part, now of another. (Harvey 1847: 335-336)

It is only during the process of generation itself that the generic matter from which the organism is formed is attracted and transformed into various specific materials required for the formation of various specific parts. This does, however, raise the following questions: from what matter is the organism formed there, and what is the cause of its transformation. To this question, Harvey gives the peculiar response that the cause of generation is the *egg*, which is "not only [...] the *matter* or that from which, but the *efficient* or that by which the pullet is engendered. In which finally no part of the future offspring exists *de facto*, but in which all parts inhere *in potentia*" (Harvey 1847: 272). This identification of the egg as the nucleus of animal generation has three consequences that are important for the current discussion.

The first consequence is that Harvey is able to generalize his findings on the basis of his observations of chicken-eggs and deer-embryos to animal generation as such. This

was far from obvious in his time, since most anatomists and medical authors were highly skeptical of comparisons between different animals, or at least between higher and lower animals (cf. Schmitt & Webster 1989: 60)<sup>3</sup>. Moreover, this generalization goes against Aristotle's central division of animals (In *History of Animals*), as far as their mode of generation is concerned, into vivipara, ovipara and vermipara, and its concomitant distinction between three products of engenderment: the egg, which the stagirite describes as "a certain completed result of conception out of which the animal that is to be develops-from a part of it at first, while the rest serves for food as it develops", the grub ("a thing out of which in its entirety the animal in its entirety develops, by differentiation and growth of the embryo"), and the living creature itself (489a-489b). It is clear, however, that Harvey nevertheless seeks to harmonize his own theory with Aristotle's (cf. Pagel 1967: 45), obviously inspired by such remarks from the *History of Animals* as: "Of viviparous animals, some hatch eggs in their own interior, as the selachia; others engender in their interior, as man and the horse" (489a). Such a peculiar violation of the distinction between these two modes of generation soon leads Harvey to claim (after abolishing, again allegedly on Aristotle's own authority; the distinction between the egg and the grub (Harvey 1847: 457-458)):

Generation in both is one and identical in kind: the origin of either is from an egg, or at least from something that by analogy is held to be so. An egg is, as already said, a conception exposed beyond the body of the parent, whence the embryo is produced; a conception is an egg remaining within the body of the parent until the foetus has acquired the requisite perfection. [...] Wherefore, the same theorems and conclusions, though they may appear paradoxical, which we drew from the history of the egg, turn out to be equally true with regard to the generation of animals generally. (Harvey 1847: 462-463)

In this way, Harvey reduces the three modes of generation to one species, which allows him to generalize or draw consequences for other cases than the one under direct examination through reasoning through analogy (Harvey 1847: 462). James Lennox (2006) has argued that this move is, in itself, perfectly in line with Aristotelian epistemology, for Aristotle regarded processes in individuals in as far as they exemplified behaviour of their kinds (species). By extending and unifying the species, Harvey is able to claim wider scope for his findings and explanations. But there is

---

<sup>3</sup> In the second response to Riolan, Harvey responds to critics who opposed his alleged "vainglorious love of vivisections, and who scoff at and deride the introduction of frogs and serpents, flies, and others of the lower animals upon the scene" (Harvey 1847: 109), and advises them: "If for the sake of studying the meaner animals you should even enter the bakehouse with Heraclitus, as related in Aristotle, I bid you approach; for neither are the immortal Gods absent here, and the great and almighty Father is sometimes most visible in His lesser, and to the eye least considerable works" (Harvey 1847: 110).

another aspect to Harvey's operation, for it not only generalizes his consequences, but also unifies the problem: there is henceforth one problem of animal generation, concerning the process of generation as such, not a set of distinct though related problems concerning distinct though related processes of generation.

A second consequence of Harvey's focus on the egg as a material and an efficient cause is that the distinction of the sexes now no longer coincides with the distinction of form and matter, for which Aristotle had provided an ingenious though highly questionable proof in *Generation of Animals*:

The female does not contribute semen to generation, but does contribute something, and [...] this is the matter of the menstrual flow, or that which is analogous to it in bloodless animals [...]. For there must needs be that which generates and that from which it generates; even if these be one, still they must be distinct in form and their essence must be different; and in those animals that have these powers separate in the two sexes the body and nature of the active and the passive sex must also differ. If, then, the male stands for the effective and active, and the female for the passive, it follows that what the female would contribute to the semen of the male would not be semen but the material for the semen to work upon. This is just what we find to be the case, for the menstrual blood has in its nature an affinity to the primitive matter. (729a)

This proof runs essentially on the purported distinction between activity and passivity and between form and matter, and the idea that in the higher animals, these *distinct* roles are really *separated*. Although Harvey maintains much of Aristotle's insights and framework, he does present the egg as something which violates the metaphysical division of labour that the latter sees at work in the more perfect products of nature. This is important because it shows that the theory of epigenesis is not necessarily committed to the division between an active (efficient) and passive (material) role in generation, and a distinctness of the form and the matter of generation, let alone the arguably sexist connotations of these principles (as has been claimed by Helmut Müller-Sievers 1997).

A third consequence of Harvey's conception of the egg is that the egg is now not only an autonomous entity, but also the nexus of a whole series of oppositions, and the locus of mediation and perpetuation:

The egg also seems to be a certain mean; not merely in so far as it is beginning and end, but as it is the common work of the two sexes and is compounded by both; containing within itself the matter and the plastic power, it has the virtue of both, by which it produces a foetus that resembles the one as well as the other. It is farther a mean between the animate and the inanimate world; for neither is it wholly endowed with life, nor is it entirely without vitality. It is still farther the mid-passage or transition stage between parents and offspring, between those who are, or were, and those who are about to be; it is the hinge and pivot upon



which the whole generation of the bird revolves. The egg is the terminus from which all fowls, male and female, have sprung, and to which all their lives tend, -it is the result which nature has proposed to herself in their being. And thus it comes that individuals in procreating their like for the sake of their species, endure for ever. The egg, I say, is a period or portion of this eternity; for it were hard to say whether an egg exists for the sake of the chick that it engenders, or the pullet exists for the sake of the egg which it is to engender. Which of these was the prior, whether with reference to time or nature, - the egg or the pullet? (Harvey 1847: 271)

Here, we meet one of the many forms of the motif of circularity that permeates Harvey's work. This motif has also been recognized in his account of pulmonary circulation, which has been traced back to Aristotle's analysis of the weather cycle (Gregory 2001). More important, however, is the underlying Aristotelian idea that "circular motion [...] is the only motion which is continuous. That, too, is why all the other things- the things, I mean which are reciprocally transformed in virtue of their qualities and powers, e.g. the simple bodies-imitate circular motion" (337a) (Cf. Pagel 1967: 83).<sup>4</sup> In Harvey, the sublunary circularities of pulmonary circulation and animal generation are such perishable analogues of eternal motions: they perpetuate what is otherwise fragile and ephemeral: living structure.

It is for this reason that Harvey regarded the circulatory system to be the first part of the body that was formed. Any other organ fulfils its specific function only within the whole of the anatomical structure, and for that very reason cannot exist in itself. The circulatory system, on the other hand, is precisely the anatomical backbone and basis of the closure that characterizes complex organization. The egg, too, fulfils its role in generation because it forms the mediation between the otherwise inexplicable opposites. Harvey thus embraces circularity as central in both his anatomy and his theory of generation: In anatomy, he argues that the animal is a semi-self-perpetuating function built around a circular, closed system, namely the circulatory system; in embryology, he tries to show that the structure of the animal is both prior and posterior to its matter, since it is both its cause and its effect.

The centrality of the assumption of circularity in Harvey has often been overlooked because we expect a solution to the problem of animal generation in terms of mechanisms, vital properties or intentional steering of the process. Since it is hard to read Harvey's account of generation as mechanist, it is natural to view it as vitalist.

---

<sup>4</sup> In Aristotle, circular movement is the movement of the fifth essence, the only natural movement that is stable and perpetual, and is therefore strongly associated with eternity. In the sublunary realm, however, there is constant generation and corruption. Nevertheless, as Aristotle himself advances in *Meteorology*, "This world necessarily has a certain continuity with the upper motions" (339a).

James Bono (1990), for instance, argues that, although Harvey rebuked renaissance theories of medicine for invoking transcendent “spirits” that “animate” the otherwise inert or insufficient biological matter, he proposed instead another kind of “vitalism” that makes activity “immanent” to biological matter. I object that, according to Harvey, it is only in the circulation that blood is vital:

The blood considered absolutely and by itself, without the veins, in so far as it is an elementary fluid, and composed of several parts-of thin and serous particles, and of thick and concrete particles called cruor-possesses but few, and these not very obvious virtues. Contained within the veins, however, inasmuch as it is an integral part of the body, and is animated, regenerative, and the immediate instrument and principal seat of the soul, inasmuch, moreover, as it seems to partake of the nature of another more divine body, and is transfused by divine animal heat, it obtains remarkable and most excellent powers, and is analogous to the essence of the stars. (Harvey 1847: 510)

This suggest that vitality is due to the closure in which life figures, not to an internal property of the matter involved.

Similarly, Harvey’s use of analogies with intentionality cannot be taken at face value. As Benjamin Goldberg has pointed out with respect to Harvey’s use of the analogy with intentional agency:

it is not an opposition between natural causation and creativity that is fundamental, but rather a mismatch between natural causation and the orderly specificity and functional complexity of generation. The sorts of efficient and material causes Harvey could rely on, heat, movement, and so on, were far from being able to account for the product of generation: such causes could not create something that happened in a regular manner (first this part, then this, in all kinds of animals observed) nor something that was made up of many individual complicated parts and yet was functionally integrated towards the well being of the animal. (Goldberg 2013: 430)

It appears that for Harvey, the circularity exhibited by living systems is not something that can be done away with by mechanist, vitalist or animist assumptions. As a result, his theory of epigenesis leads us to embrace circularity in accounting for animal generation.

### **2.2.2 The Disunity of Preformationism**

Harvey’s attempt to revive Aristotelian epigenesis did not find much following in the following half of the 17<sup>th</sup> century. The explanation for this has traditionally been Harvey’s overdependence on Aristotelianism, which could not charm the many

philosophers and men of medicine who had started to embrace the mechanical philosophy, and the flimsiness of his observational resources due to his reliance on his own eyes and single lenses. Buffon's (1749b: 118-119) initial judgment that Harvey was unfortunate in not having a microscope at his disposal, therefore having to rely on images of extremely low resolution in order to find out the minute origins of biological form, was followed not only by early 20<sup>th</sup> century historians of embryology such as Francis Cole (1930: 143-144) and Joseph Needham (1959: 148-149), but equally by contemporary philosophers (Press & Tanur 2001: 73) and historians (Goldber 2013: 427) of science. But this judgment may say less of the failings of Harvey's approach than of our willingness to get carried away by the rhetoric surrounding early modern microscopy - which was immense, as is evident from the preface to Robert Hooke's both scientifically and commercially successful *Micrographia*:

By the additions of such *artificial Instruments* and *methods*, there may be, in some manner, a reparation made for the mischiefs, and imperfection, mankind has drawn upon it self [...] The next care to be taken in respect of the Senses, us a supplying of their infirmities with *Instruments*, and, as it were, the adding of *artificial Organs* to the *natural*; this in one of them has been of late years accomplish with prodigious benefit to all sorts of useful knowledge, by the invention of Optical Glasses. By the means of *Telescopes*, there is nothing so *far distant* but may be represented to our view; and by the help of *Microscopes*, there is nothing so *small*, as to escape our inquiry; hence there is a new visible World discovered to the understanding. (Hooke 1667)

Of course, Hooke does not tarry to confess which philosophical enterprise and world view is supported by this new regimentation and expansion of the senses:

It seems not improbable, but that by these helps the subtilty of the composition of Bodies, the structure of their parts, the various texture of their matter, the instruments and manner of their inward motions, and all the other possible appearances of things, may come to be more fully discovered; all which the antient *Peripateticks* were content to comprehend in two genera land (unless further explain'd) useless words of *Matter* and *Form*. From whence there may arise many admirable advantages towards the increase of the *Operative*, and the *Mechanick* Knowledge, to which this Age seems to much inclined, because we may perhaps be inabled to discern all the secret workings of Nature, almost in the same manner as we do those that are the productions of Art, and are manag'd by Wheels, and Engines, and Springs, that were devised by humane Wit. (Hooke 1667)

Jordynn Jack (2009) has pointed out that the language of the *Micrographia* follows a rhetoric that is intended to aid both the readers and other microscopists to figure out what exactly they are seeing through the lense of the microscope. On the one hand, Hooke systematically suggested that microscopic bodies were *analogous to macroscopic*

bodies, and on the other he presented them as *mechanical in structure*. The importance of this is that it facilitated the spread of the idea that the microscope reveals worlds of similar levels of complexity at lower levels of magnitude, and that it unveils the mechanical structure and substructure of reality.

Both of these ideas are clearly present in the early microscopists responsible for the alleged evidential basis of the theory of preformation, namely Jan Swammerdam, Marcello Malpighi, Antoni van Leeuwenhoek and Nicolaas Hartsoeker. Swammerdam's main contribution is twofold, since he delivered both an argument against the distinction between the generation of perfect and imperfect animals, and a paradigm for the "unfolding" of pre-formed structure, the chrysalis. Whereas Harvey had denied that metamorphosis could account for the higher animals, but still believed it to be possible for lower animals, Swammerdam revolted against this distinction between higher and lower animals as a distinction between levels of complexity, and therefore ruled out the metamorphosis and spontaneous generation of even the lower animals (Swammerdam 1737: 2) He instead argued that we can already find, in the chrysalis, the full-formed butterfly, and that it becomes the butterfly not through any change of its limbs, but through mere growth of its pre-existent parts (Cf. Swammerdam 1737: 6-7).

Marcello Malpighi too gave an important impetus to the development of preformationist embryology with his 1673 *Dissertatio epistolica de formatione pulli in ovo*. In this dissertation, addressed to the Royal Society of England, Malpighi meant to advance on Harvey's discoveries by offering his own, microscope-assisted, observations. The most fateful of his corrections is his claim that a relatively articulated form of the fetus can be found in the egg even prior to incubation. This was taken to disconfirm Harvey's view that the embryo started from very simple beginnings, and only later acquired the features now seen in it from the start (Malpighi 1687: 53-54). Besides this alleged observational evidence, we find in Malpighi's work a clear framework in which to fit it, namely that of a new development in mechanicism where the living being is seen as a conglomerate of an enormous amount of minute machines, as a mechanism consisting of mechanisms (Malpighi 1697: 104; cf. also Giglioni 1997: 162-163).

Whereas Malpighi's observations suggested that the animal somehow preexists in the ovum, two Dutch microscopists were conducting microscopical observations of animal seminal fluid. In spite of a claim to originality by Hartsoeker, it is mostly the name of Leeuwenhoek that remained associated with the discovery of the animalcules. Over the course of several years, Leeuwenhoek addressed to the Royal Society several letters in which he claimed to have seen, in the seminal fluid of male animals, an enormous amount of tiny animals, "animalcula" (*Dierkens*), each of which seemed to move of its own accord. He maintained that these animalcules could not have been formed by spontaneous generation, nor could they be inorganic in nature, given that they "are made out of so great a multitude of parts as [...] the multitude of parts out of which our bodies are composed" (Leeuwenhoek 1686: 10; my translation).

For many readers today, it seems obvious that the reports of the early microscopists do not yield the preformationist theory as we now know it. For this, two other elements were required, as Margaret Wilson has seen. The first is that there were already anticipations of these discoveries in the speculations of Gassendi and Descartes, who may not themselves have endorsed preformation, but in any case prepared its theoretical discourse (Wilson 1995: 121). The second is that it required the theoretical perspectives of the likes of Malebranche and Leibniz to make preformationism into a coherent theory (Wilson 1995: 128).

It is telling, first of all, that Malebranche's famous discussion of preformation appeared first in the course of a discussion of the reliability of the senses in the quest for truth. Our senses, he tells us, are designed for our practical dealings, our direct utility, not for theoretical knowledge, as the newest developments in observation, i.e. the turn to the microscope, reveal. What appears to the naked eye as homogeneous, and sterile, the microscope shows to be heterogeneous and teeming with life. Our own perception, focused as it is on our own order of magnitude, is oblivious to the complexity at other levels.

In offering the theory of preformation as an illustration of the errors we may incur when we fail to assess the limitations of the senses, Malebranche also seems to be indicating that he was aware of the gap between the observational data offered by early microscopy and the theory of preformation and preexistence. That may be why he argues that reason allows us to use imagination in extending sense beyond its current boundaries. He believes this to be warranted by his doctrine that space is infinitely divisible and hence does not bottom out mereologically (OC I: 81). Since we know from observation that below the order of magnitude revealed to our senses lie several orders of magnitude that are equally complex and structured, and since reason tells us that there can be no lower bound to orders of magnitude, there is no reason to suppose that complexity decreases along with the magnitude of things. This argument provides Malebranche with a possible justification of what many have since taken to be an outrageous fiction, namely that each organism came into the world fully formed at creation, was embedded in its parent (in *casu* the mother) and from then on merely awaits the process of growth through orders of magnitude until it reaches our own size:

Nor does it seem unreasonable to believe even that there is an infinite number of trees in a single seed, since it contains not only the tree of which it is the seed but also a great number of other seeds that might contain other trees and other seeds, which will perhaps have on an incomprehensibly small scale other trees and other seeds and so to infinity. So that according to this view, which will appear strange and incongruous only to those who measure the marvels of God's infinite power by the ideas of sense and imagination, it might be said that in a single apple seed there are apple trees, apples, and apple seeds, standing in the proportion of a fully grown tree to the tree in its seed, for an infinite, or nearly infinite number of

centuries; that nature's role is only to unfold these tiny trees by providing perceptible growth for that outside its seed, and imperceptible yet very real growth in proportion to their size, for those thought to be in their seed...

Likewise, a chicken that is perhaps entirely formed is seen in the seed of a fresh egg that has not been hatched. Frogs are to be seen in frogs' eggs, and still other animals will be seen in their seed when we have sufficient skill and experience to discover them. But the mind need not stop with the eyes, for the mind's vision is much more extensive than the body's. We ought to accept, in addition, that the body of every man and beast born till the end of time was perhaps produced at the creation of the world. My thought is that the females of the original animals may have been created along with all those of the same species that they have begotten and that are to be begotten in the future. (OC I: 82-83; Pyle 2003: 169)

Like Malebranche, Leibniz too referred approvingly to the findings of the microscopists as an occasion for rethinking the nature of organic form:

this is where the *transformations* of Swammerdam, Malpighi, and Leeuwenhoek, the best observers of our time, have come to my aid, and have made it easier for me to admit that animals and all other organized substances have no beginning, although we think they do, and that their apparent generation is only a development, a kind of augmentation. I have also noticed that the author of the *Search after Truth* [i.e. Malebranche], Régis, Hartsoeker and other able persons have held opinions not far removed from this. (G IV: 480; Leibniz 1989: 140)

Justin Smith (2011) has argued that the encounter with the early microscopists was a major cause of the shift in Leibniz's thinking on living form from the macroscopic level to the microscopic, and the concomitant shift from a conception of an organic body as a hydraulico-pneumatico-pyrotechnical machine to a nested individual composed of other biological individuals (Smith 2011: Chapter 5). He has also advanced that we should therefore take the doctrine of preformation expressed throughout Leibniz's published and unpublished writings alike to be more central to his metaphysics than has traditionally been conceded. Most importantly, he suggests that it reveals a close analogy between preformation and preestablished harmony. Leibniz sometimes presents the theory of preformation as the embryological analogue of the doctrine of preestablished harmony. Smith (2011: 194-195) has argued that this may have been connected with some more philosophical and theological reasons Leeuwenhoek adduced for embracing preformation.

However, it has been remarked that Leibniz's ideas cannot be based entirely on the microscopic discoveries to which he alludes, since these hardly *show* the infinite composition of extension. Like Malebranche, he needs a further element to go from that which can reasonably be taken to be the observational data (the observation that organisms do not bottom out in the macroscopic), to the idea they are meant to ground (the idea that organisms do not bottom out at all). The additional element is, as Ohad

Nachtomy (2014) has convincingly argued, that of the actual infinite. Leibniz is committed to the idea that everything in nature is mechanical, but that some mechanically structured entities are of a complexity that transcends our capacity of understanding absolutely: they are divine machines. He thus seems to be committed to the idea that organic bodies do not differ qualitatively from artefacts, *and* to the idea that they differ not merely in degree from artefacts. These two seemingly contradictory ideas are brought together by the assumption that organisms are machines within machines to infinity, and therefore infinitely complex, whereas artefacts bottom out in a lowest level of relevant mechanical structure:

Thus each organized body of a living being is a kind of divine machine or natural automaton, which infinitely surpasses all artificial automata. For a machine constructed by man's art is not a machine in each of its parts. For example, the tooth of a brass wheel has parts or fragments which, for us, are no longer artificial things, and no longer have any marks to indicate the machine for whose use the wheel was intended. But natural machines, that is, living bodies, are still machines in their least parts, to infinity. That is the difference between nature and art, that is, between divine art and our art. (G VI: 618; Leibniz 1989: 221)

The distinction is not one of degree because, for the rationalists at leasts, the infinite may be quantitative, but it is radically distinct from any determinate quantity (it is at once somehow situated in the progression of the series of natural numbers and yet not any member of that series). They thus need the notion of a positive infinite to arrive at a mechanical theory of organic complexity. This reveals that, in thinking of the structure of life, they too admit a regression, but that they and higher than every member of that series. Like Malebranche, Leibniz needs to have replaced Harvey's circularities with a linear regression of machines within machines within machines...

Another problem faced once we accept Smith's general reading is that of the relation between occasionalism and pre-established harmony. Smith argues that Leibniz would agree with the general idea that preformation is the better theory because it has the same benefits that the theory of preestablished harmony has. After all, if one agrees that organisms cannot come into existence through mere mechanical processes, than one would have to believe that the coming to be of a new organism would require a miraculous intervention. The latter position, however, is analogous to what Leibniz made of the occasionalist position, namely the requirement of constant divine miraculous intervention in the natural order. It seems better, therefore, to assume that God made all the organisms that would ever exist in the world at the moment of creation, which then go on to develop through merely mechanical means. This is analogous to the theory of preestablished harmony, according to which God made all substances at the moment of creation and then let these internally develop according to the general principles of nature.

It is interesting to note here that the parallel between theories of causation and mind-body interaction on the hand and theories of generation is best known from the works of Kant. As we shall see, Kant will refer to it on two important occasions of his career. This has led Karen Detlefsen (2003) to inquire into the reason for the peculiarity that Malebranche embraced both preformation and occasionalism, and to conclude that the difference resides in Malebranche's and Leibniz's conceptions of occasionalism. Leibniz regarded generation in nature as only possible through an immediate particular volition of God, namely a supernatural, miraculous intervention in the regular order of nature, and occasionalism as the doctrine that all causal interactions are of this kind. This analogy is likely to be false, since Malebranche's occasionalism does not imply that all causal interactions are exceptional, supernatural interventions, but merely that all causal powers ultimately reside in God, and that causal interactions depend on the powers and volitions of God, be they general or particular. On such a reading, many distinctive differences between occasionalism and preestablished harmony start to blur. As we will see in subsection 2.3.1, Kant will use this fading boundary to argue against preformationism's claim to theoretical and doctrinal superiority.

A third implication of Smith's analysis is that it renders inoperative, to a certain extent, the traditional distinction between two forms of preformation, namely animalculism and ovism. Smith (2011: 183-184) notes that Leibniz seems to have preferred the spermist or animalculist alternative, but agrees with e.g. Pinto-Correia (1997: 93) that Leibniz ultimately never made a final decision between the two alternatives. This distinction was important, however, not only because it reflects the decision on which sex holds the key to procreation and on similar ethico-theological questions, but also because they lead to preferences between the two distinct versions of preformationism listed by Charles Bonnet in his 1762 *Considérations sur les corps organisés*: "emboîtement" and "dissémination". The first theory is that all generations are embedded in one another, whereas the latter is that germs are disseminated throughout the whole of nature and develop only once they have found a suitable body in which to develop (Bonnet 1779: 2-3). Bonnet decides in favour of neither theory, but the fact that he notes the distinction may already be useful to contrast the theories of Malebranche and Leibniz. Whereas the former seems to prefer an account where our genealogical lineage is traced by the individuals in which we have been mereologically embedded for all time, the latter sees biological individuals as independent from this and capable of entering into all kinds of organic composites throughout their histories.

In order to properly understand preformationism we must dispel one final misunderstanding, namely the persistent idea that the preformed individual is a mere miniature, however minuscule, of the full-grown form. Almost none of the preformationist theorists I have discussed here hold this view; in fact, almost all of them stressed the great degree of complexity of the process of unfolding required to bring out the eventual form and maintained that, although the preformed structures were of such



a kind as to be empirically almost unrecognizable as the adumbrations of their adult forms, they did contain all the necessary elements for these forms, and could yield them through mechanical means alone. Preformationist theories could thus differ on 1) the degree of surface similarity between the preformed and the developed form, and 2) the means admissible in our explanations of how these adult forms come about. This flexibility constituted one of the great strengths of preformationist theory, but may also ultimately have led to its downfall as it was stretched beyond its usefulness as an umbrella-term.

### 2.2.3 Metamorphogenesis

In the literature on early modern theories of generation, it is customary to treat the debates raging in that era as between the two rival theories of epigenesis and preformation, and to categorize all other theories as peculiar variants of the former two. I believe this to be mistaken, since it overlooks the extent to which both alternatives reacted against an older alternative that continued to resurface over the course of history: the theory of metamorphosis. I have already pointed out that both Harvey and Swammerdam saw themselves as reacting against theories that understood generation as spontaneous generation or a process similar to it. In this subsection, I will argue that the theory of metamorphosis and its associated theory of spontaneous generation is broader than what we now understand under it, and that it captures elements of what we have come to see as borderline cases of preformation and epigenesis. Allowing for a third category makes better sense of what scholars have previously tried to grasp through awkward distinctions between forms of preformation or epigenesis.

Many authors have found it necessary to distinguish between mechanical and non-mechanical theories of epigenesis. Mechanical theories describe the process of epigenesis as involving only the matter and the processes allowed for by classical mechanicism. They are contrasted with animist or vitalist theories of epigenesis, which regard the process of epigenesis as involving vital matter and/or vital forces. The problem with this distinction is that it usually forces us into allowing that the more mechanical brands of epigenesis assume a greater degree of preformation than the more vitalist variations. This is then often accounted for by distinguishing, within preformation, between preformationism proper, the theory that the structure is formed before conception, and preexistence, the theory that the structure is not even formed by the genitor before conception, but by God at moment of the creation of the universe (Roger 1963: 326; Bowler 1973: 259). This allows for a third position, according to which the structure is already very heterogeneous at the moment of conception. Once this heterogeneous structure is formed relatively unintentional forces suffice to finish the work of generation. I will contend here that a variety of theories that have been

classified as “preformation but not preexistence” and “mechanical epigenesis” are in fact a species of the oldest dominant theory in the tradition of embryology, and stems from the works of Hippocrates and Galen. Further on in this chapter, I will show that Kant was attentive to this distinction, although we, his readers, are usually not, and that this leads to misinterpretations.

The oldest version of the theory of metamorphosis, as I will call it here, following Harvey, is that of Hippocrates. It is also known as the theory of the two semina, since it assumes that both parents contribute seminal material to the foetus, and that this seminal material carries information about the body of the parent. For Hippocrates (2012: 6-23), the surplus of humours used by the different organs is sent to the reproductive organs as seminal material, and is composed of the correct mixtures required to (re)constitute those organs. Once the two semina are brought together, the materials rapidly align and cohere in such a manner as to form the first rudiments of the embryo. Since the semina contain the material of all the organs, however, these first rudiments are already a very highly articulated structure. Mind that this structure is, in an important way, already present in the semina, because it already contains all the materials required to make the foetus. Metamorphosis theories usually face the challenge of explaining why, in higher animals, spontaneous generation only occurs after conception, and not within the reproductive organs of the parents, given that these already contain all the required elements.

The Galenic theory is related, but it is important to note how it employs the idea of a natural faculty. For Galen, a central problem of physiology and anatomy is that of explaining how organs and organisms attract the specific foodstuffs they need and adapt them to their own structures. He rails against the systems of two rivals, which argue that only mechanical processes need be involved in these processes. They assume, for instance, that the foodstuffs are corpuscles of a specific kind and that the tissues of the organs contain pores suited to these and only these corpuscles. Galen (1916: 43-49) finds such accounts wanting since they deny rather than explain the selectiveness of living tissues. He postulates in organs non-mechanical properties that select, attract, adapt and employ their nutrients.<sup>5</sup> On his account, we pass on to our offspring both the specific mixtures we have prepared and the faculties required to maintain and augment them(1916: 19-27).

---

<sup>5</sup> It is unclear whether these natural faculties are well understood as intentional-teleological. I believe they are not, given the care with which Galen distinguishes them from the faculties of the soul. Galen criticizes Aristotle for glossing over the differences between the natural faculties associated with organic structure or tissue, and the psychological faculties associated with a soul. He feels no need to assume any such thing as a vegetative soul, and therefore the associated functions cannot truly be understood as soul-like.

This theory remained dominant until the sixteenth century in a variety of versions. The most important one is arguably the canonical version presented by Fernel in his monumental *Physiologia*. It is important not just because it was highly influential, but also because it was one of the primary targets of the early mechanical philosophies like that of Descartes. Descartes probably knew of Fernel through his endorsement by the Jesuit scholastic textbooks that were a major ingredient of his intellectual formation at La Flèche. For this reason, one could argue that what Descartes offered us in his physiology, and specifically in his theory of generation, was an attempt to salvage the corpuscularian-mechanicist positions criticized by Galen by offering a mechanistic account of the Galenic theory. In doing so, he did not really offer a new theory, but rather a new foundation for an old one. This picture fits neatly into the view of Descartes's physiology advanced by Thomas Hall (1970: 64): "what Descartes had to offer were not explanations of fact. They were explanations, rather, of other peoples' explanations (often dismembered and reassembled with various additions and deletions)".

Descartes's theory of generation has been regarded by many, perhaps even by Descartes himself (Cf. Des Chene 2001: 32) as overly sketchy and implausible. It seeks to derive the form of an organic body from the mechanical interactions of the corpuscles of fluids without admixture of any teleological or intentional aspects that allow for the proper differentiation and allocation of parts. This is undoubtedly why many have thought Descartes's model to be epigenesist: it shows how structure is articulated through the differentiation of homogeneous matter. Nevertheless, there are reasons to doubt such an interpretation; after all, Descartes does need to locate the reasons why the seminal fluid gives rise to the specific structure it yields somewhere. The hint may come from a comment made by Dennis Des Chene in his presentation of the Cartesian account of generation:

A particle of blood is merely a bounded region of stuff. The only "information" it carries is its *present* quantity of motion and the determination or direction of that motion. These two properties entirely determine the outcome of any collision the particle enters into; collisions, moreover, are the only interactions of particles. Particles therefore exhibit neither memory nor foresight. (Des Chene 2001: 41)

What Des Chene reveals in this passage is the only place where Descartes can locate the information that will yield the future structure: the specific motions and directions of these motions in the seminal fluid. In this way, he repeats the problem I addressed in the previous chapter concerning his cosmogony: he cannot account for the specific structure of the end-process of generation except by somehow locating its full structure in the initial state. By refusing any finality to generative processes and any non-mechanical properties to matter, he is forced to admit either that the initial structure fully determines the outcome through its mechanical features alone or that his theory

fails to account for the specificity exhibited by living processes. For a long time, Cartesians assumed that Descartes succumbed to the latter failure and that it was best remediated by simply admitting that the initial state contains all the required information, taking this view to its preformationist extreme. It is only in the eighteenth century that natural philosophers inspired by Newton would once more begin to seriously consider alternatives.

The two main figures associated with the revival of metamorphosis in the early to mid-eighteenth century are Maupertuis and Buffon. I have already discussed their importance for Kant in the first chapter of this dissertation. Here, I will focus on their fateful interventions in the debate on animal generation and their attacks on the then-dominant theory of preformation. Both authors have been read as epigenesist, with the debate ultimately turning on the degree of their adherence to some version of preformation without preexistence, and to their preference for either mechanicism or some form of vitalism. I will argue that Buffon and the early Maupertuis are best read as seeking to provide, like Descartes, a mechanical theory of metamorphosis. Unlike Descartes, however, they believe themselves to be entitled to a wider variety of basic properties in mechanical explanation because of their Newtonianism. One of the major benefits of radical Newtonianism is that it allows only empirical research and nature to dictate which the basic properties of matter and motion are in terms of which we should explain everything natural. Much like Newton showed that we could not account for even the most basic mechanical interactions without assuming gravity or solidity, however metaphysically problematic these notions may be to a die-hard mechanist, so too Buffon attempted to demonstrate the indispensability of his own theory for any account of the phenomena of life:

I have admitted in my explanation of development and reproduction, first of all the received mechanical principles, then those of the penetrating force of gravity that we are obliged to admit, and through analogy I believed myself capable of saying that there are other penetrating forces at work in organized bodies, as experience assures us. I have proved through facts that matter tends to organize itself, and that there are an infinite number of organic parts. I have thus done nothing but generalize the observations, without advancing anything at odds with the mechanical principles, if one understands by that word what we should understand by it, i.e., the general effects of Nature. (Buffon 1749b: 52-53; my translation)

From this starting point, Buffon's theory differs greatly from that of his preformationist predecessors. He denies the infinite divisibility of matter, and thereby prohibits recourse to the infinite envelopment as an account of the complexity of organisms (Buffon 1749b: 27). Instead, he advances that we have been mistaken in believing that because the world bottoms out, it bottoms out in particles of a geometrically simple shape (typically spheres and its eccentric variants or platonic polyhedra) (Buffon 1749b:

22). If we let go of this bias, we can hypothesize that the world contains an infinite variety of organic molecules, which are the rule in nature rather than the exception, and which can produce the less complex ones through degenerations (Buffon 1749b: 39). The same bias towards traditional mechanical properties in terms of extension has caused us to overlook the idea of a *moule intérieure*:

In the same way that we can make molds by means of which we give to the exterior of a body the shape that we please, let us suppose that nature can make molds through which she gives, not only external shape, but also the internal form; would this be a way through which reproduction could operate? (Buffon 1747b: 34; my translation)

The concept of an interior mold allows Buffon to set in place his nutritional theory of generation, which is clearly reminiscent of the Hippocratic-Galenic theory:

In order to properly understand this manner of reproduction, it suffices to picture that in the nutriments that these organized beings attract, there are organic molecules of different species, that by a force similar to that which produces gravity, these organical molecules penetrate all the parts of the organized body, which produces the development and makes the nutrition, that each part of the organized body, each interior mold admits only the organical molecules that are proper to it, and finally that when development and growth are almost finished, the surplus of organical molecules that served it previously are sent from each part of the individual to one or more places, where gathered together they form through their union one or several little organized bodies, that have to be completely similar to the first individual, since each of the parts of this individual has sent the organic molecules that are most analogous to it, those which would have served in its development, if it had not been completed, those which through their similarity could serve for nutrition, those finally that have nearly the same organic form as the parts themselves. In such a way, in all species where a single individual produces its like, it is easy to draw the explanation of reproduction from that of development and nutrition. (Buffon 1749b: 54; my translation)

Like in the Hippocratic-Galenic theory, the foetus is formed through the rapid association and alignment of thoroughly prepared materials that are direct products of the organs of the parents. On such an account, organisms could also spontaneously generate out of present organic materials, even though this would be highly unlikely. Again, the theory of metamorphosis seeks to fend off epigenesis by assuming a mitigated form of preformation without preexistence, and again it needs to dissociate itself from preformation by exaggerating the specificity of the processes involved in generation.

A similar account is to be found in Maupertuis' somewhat earlier theory. In the earlier phases of his thought, Maupertuis, like Buffon, believed that by adding to the

restricted list of mechanical forces the broader but not terribly exotic list of chemical forces, we could account for the specificity and selectivity involved in the phenomena of life without having to reinstate Galenic faculties (Maupertuis 1768b: 88-89). In his case, however, this optimism was more short-lived: later in his life, he admitted that not even the extended list of mechanical properties could account for the phenomena of interest to him: generation and heredity (Cf. Roger 1963: 483). It is at that point that he introduced a notion of material particles that could have memory of states in which they had previously found themselves and desires for and aversion towards specific other (kinds of) particles (Maupertuis 1768b: 146-147). Throughout his career then, Maupertuis moved from a Newtonian extended mechanist position to a position best described as “vital materialist”<sup>6</sup>. We will see that, for Kant, these two options were always dangerously close to each other in any case.

The position I called metamorphosis amounts to the following: the idea that the specificity of organic form is due to a precipitate association of a thoroughly prepared material through mechanical or quasi-mechanical processes. The difference between this position and that of either epigenesis or preformation has often been observed and taken to be puzzle. McLaughlin (1990, 20), for instance, contrasts *epigenesis* with *pangeneses*, the latter being the Hippocratic theory I just sketched. Bowler (1971: 223) adopts Harvey’s term, like I have done here. Joseph Needham (1959, 183-184) characterizes it as precipitation followed by preformation. All these names reveal the peculiar position of this theory amidst the otherwise clean-cut opposition between preformation and epigenesis.

#### 2.2.4 Epigenesis and the observability of structure and genesis

As a conclusion to this section, I will briefly discuss the revival of epigenesis in the hands of Caspar Friedrich Wolff and Johann Friedrich Blumenbach. My intention here is not to give detailed accounts of their theories, but to indicate the issues important for my presentation of Kant’s theory. The first issue is that of the status given to observation in debates over preformation, and the second that concerning the guiding principle of development. These two themes resonate with the main issues involved in Harvey’s account, leading us to the updated version of the latter’s theory of epigenesis.

---

<sup>6</sup> This term was advanced by Lenoir to denote a position originally advanced by Friedrich Blumenbach and diffused through German scientific culture through the approval of Immanuel Kant (Cf. Lenoir 1982). The position is thoroughly materialist, but ascribes to matter an activity and a vitality that traditional mechanicism denied to it. More recently, Charles Wolfe (2011) has suggested that a similar position was endorsed by Enlightenment Materialists like Diderot and La Mettrie.

In my discussion of preformationism, I already indicated that this theory required its proponents to go beyond the observational data with which microscopy had furnished them. As microscopes, or at least researchers' skill with them, improved, no advance was made on these discoveries. On the contrary: the idea of straightforward preformation became less plausible as the rudiments of organisms more and more turned out to be different from the adult structures. For Buffon and Maupertuis, this suggested that we need to return to metamorphosis, according to which it is not the structure that is present in the rudiments, but the elements, which are then assembled through natural processes. But to others, it suggested that we had given up on Harvey's theory of epigenesis too soon. One such figure was Caspar Friedrich Wolff, who dared react against the leading consensus by advancing an epigenesist position, and thereby attracting the criticism of the leading Swiss physiologist and recent convert to preformation Albrecht von Haller. Wolff's position was primarily directed at version of preformation advanced by Haller, who held that

the essential parts of the foetus are found fully formed at all times; not, of course, such as they appear in the adult animal: they are structured in a such a manner that certain and prepared causes, precipitating the growths of some of these parts, impeding that of others, changing their relative positions, rendering visible organs once diaphanous, giving consistency to fluids and to mucosity, eventually form an animal very different from the embryo, and in which there is nevertheless not a single part that did not exist essentially in the embryo. (Haller 1758: 186; my translation)

The specific feature of Haller's account is that it explained why the rudiments of the embryo could be markedly different from the full form, and more importantly, why earlier stages of the embryo could appear less articulated and therefore more homogeneous to the observer. For Haller, this is due to several parts of the organism being diaphanous and fluid at first, and acquiring solidity only later. It is this insight which lay behind his objections to Wolff's theory. Indeed, as Wolff himself would remark, Haller objected to him the usage of the axiom that "what I cannot see, is not there"(Wolff 1764: 74) – that is to say, there are no such things as invisible parts. Wolff of course denied that he ever took recourse to such a maxim, and merely stressed that his refusal to admit the existence of these preformed parts was due to his being incapable of ascertaining their presence by any means (not just visible means, but through any kind of intervention) and squaring their existence with any of the phenomena that he was able to ascertain. As far as the fluidity is concerned, Wolff remarks that the property of fluidity is opposed to the structural features of an organism (the precise interrelation of its parts) that preformation is meant to account for (Wolff 1764: 133-134).

There may be, however, a further disagreement that divides the two authors, namely their assumptions on the meaning of generation as differentiation. For Wolff, generation is a process whereby an initially largely homogeneous material is differentiated into different kinds of materials, which are simultaneously or subsequently employed in the formation of organs. For Haller, the very idea that one can differentiate what is not already differentiated may have been ludicrous. A similar assumption lay behind the metamorphosis theories of Buffon and the early Maupertuis, namely that all the requisite materials for the organic form must be present at the start in the semina, and that these are just sorted out and arranged by the forces working on them. Such an assumption is not outrageous; it lay behind the great demonstration of Newton's Opticks as well:

Newton was an atomist who believed that matter is composed of hard and permanent particles which are endowed with various properties from the beginning of creation. The properties of natural things, on this conception, depend on the connate properties of the ultimate particles which go into their constitution. Natural changes consist, not in the creation or annulment of properties, but in the separating and combining of the particles. These operations merely result in making manifest or concealing the original properties. From this point of view the dispersion of light by a prism simply *separates* the mixed corpuscles and, as a result, colours appear. (Sabra 1967: 296)

It is this assumption that epigenesists cannot share: they must assume that nature is truly creative, in that it can bring about new properties that are not actually (though they may be contained virtually) in the initial material. The processes responsible for generation cannot therefore be mere sorting out of the material. The question then becomes what in nature can effectuate such processes. For the major proponents of epigenesis in the second half of the eighteenth century, the fact of genuine formation requires us to assume the existence of formative forces. For Wolff, this is a *vis essentialis* responsible for guiding the process of formation.

The major figure we now associate with late eighteenth century epigenesis is, perhaps to no small extent due to the efforts of Timothy Lenoir (1982), Johann Friedrich Blumenbach, whose theory is perhaps expressed in the following passage from his *institutions of physiology*:

The *matter* of which organized bodies, and therefore the human frame, is composed, differs from all other matter in this, that it alone is subject to the influence of the vital powers.

Among the orders of vital powers, one is eminently remarkable and the least disputable of all,-which, while it acts upon that matter hitherto shapeless but mature, imparts to it a *form* regular and definite, although varying according to the particular nature of the matter. To distinguish this vital power from the rest, permit us to designate it by the term,-*nisus formativus*.



The *nisus formativus* occurs to the genital *matter*, when this is mature and committed to the uterus in a proper condition and under proper circumstances, lays in it the rudiments of conception, and gradually *forms* organs fitted for particular purposes; preserves this structure during life, by nourishing the body; and reproduces, as far as it can, any part accidentally mutilated. (Blumenbach 1817: 335)

The specificity of Blumenbach's position lies in its insistence on the tie between the vital powers and the matter on which they operate. As Lenoir has expressed it:

In spite of his decision to support the epigenetic theory, Blumenbach did not want to abandon what he considered to be desirable features of the preformationist account. In particular, he wanted to retain the idea fundamental to Haller's preformationism, that the fact of biological organization could not be accounted for in terms of physico-mechanical causes, but had to be treated as primary. On the other hand, he wanted to avoid treating biological organization as a result of the super-imposition upon inorganic matter of a separate force, a *Lebenskraft*, or a soul, which exists independently of a material substrate. In attempting to steer a course between the Scylla of materialistic reductionism and the Charybdis of vitalism, Blumenbach treated the agent responsible for organic structure as a kind of Newtonian force for the biological realm, which he called the *Bildungstrieb*. (Lenoir 1982: 20)

Like so many eighteenth-century life-scientists, and not coincidentally like Buffon and Maupertuis before him, Blumenbach took his cue from Newtonianism in order to postulate this special and specific force and that for him

*the word* Formative Drive [*Bildungstrieb*], to the same extent as *the words* attraction, gravity, etc. is to serve as nothing more or nothing less than to denote a certain force, the constant effect of which can be recognized in experience, but the *cause* of which is, for us, to the same extent as that of the aforementioned, however universally recognized, forces of nature, a *qualitas occulta*. (Blumenbach 1789: 25-26; my translation)

As we will see, Kant would come to appreciate this position, although I fear we may have misunderstood his reasons for doing so. What is clear at this point, however, is that Blumenbach's version of epigenesis did not resolve the issues around the theory of epigenesis, but merely brought them to the surface. Firstly, he made it clear that any valuable theory of epigenesis must postulate a teleological principle guiding the process of generation. Secondly, and more importantly, he revealed that however homogeneous epigenesis might take the initial matter on which generation draws to be, it needs to regard it as being sufficiently specific in order for it to allow for these processes of generation at all. It is the latter constraint that proved to be the biggest puzzle, and that

many now see as Blumenbach's fateful concession to preformationism – a concession believed to be made by Kant as well.

## 2.3 Epigenesis in the pre-critical period

In the previous section, I have offered a lengthy historical overview of the early modern debate on animal generation. I have done so because I fear that in relying on the historical categories as we have inherited them from the history of science, we are bound to misunderstand the specific difficulties Kant intended to tackle in his own ventures into biological theory. For a long time, Kant has been taken to be a proponent of epigenesis because of his explicit adoption of the term to describe his own biological views and in analogies meant to clarify some of his positions in epistemology, philosophy of history (of philosophy) etc... Recently, however, a shift occurred in our reading of Kant. In 2002, Philip Sloan argued that Kant in fact sought to develop a more mitigated preformationist theory throughout his career, and that only in the mid-1780s did the term epigenesis find its way into his work. When it did, however, it did not take the form of the kind of unrestricted epigenesis that had begun to take hold of German Early Romanticism. Sloan's reading was later endorsed by John Zammito (2003), and soon became the received opinion among Kant scholars.

The merit of Sloan and Zammito is certainly to have shown that we need to be careful when speaking of Kant's alleged epigenesis. I believe that they are wrong, however, in their identification of the kind of view that Kant ultimately adopted, and more importantly of the innatism they end up ascribing to the transcendental philosophy. In fact, their view comes down to the idea that Kant's generic preformationism is more preformation than epigenesis, and that Kant never truly embraced epigenesis as a theory of development. Against this, I will argue that Kant's view is a kind of epigenesis, or at least more of an epigenesist theory than the current view admits it to be, although it is meant also to address the difficulties with epigenesis.

### 2.3.1 The Only Possible Account of Animal Generation

Kant's earliest explicit dealings with embryology were in his 1763 book *The Only Possible Argument in Support of a Demonstration of the Existence of God*. This work is peculiar because of the several goals it sets itself. It starts off by criticizing traditional ontological proofs for the existence of God on the ground that existence is not a predicate, and then goes on to offer a proof of the existence of God of Kant's own making. This technical proof,

however, interesting, does not concern us here, but it is important to see that many of the discussions Kant offers in this work are attempts to show the usefulness or uselessness of certain theological and/or teleological reasonings in natural philosophy. Most importantly, Kant claims that his proof of God, or the God whose existence he proved, allows us to assume that nature constitutes a harmonious and unified order. This means, amongst others, that in nature, a great diversity of effects must be regarded as due to a limited number of principles. Kant uses this argument not only to indicate that nature would be less perfect if it required frequent miraculous intervention (AA II: 108-109) (an argument that we also saw in the *Universal Natural History* in Chapter 1), but also that we should be hesitant towards postulating further explanatory principles:

In the procedure of purified philosophy there prevails a rule which, even if it is not formally stated, is nonetheless always observed in practice. The rule maintains that in investigating the causes of certain effects one must pay careful attention to maintaining the unity of nature as far as possible. In other words, the rule maintains that one must derive a variety of effects from a single cause which is already known, and not immediately suppose the existence of new and diverse operative causes to explain different effects because of some seemingly important dissimilarity between them. Accordingly, it is presumed that there exists a great unity in nature, in respect of the adequacy of a single cause to account for many different kinds of consequences. [...] And if one finds oneself constrained to postulate a new principle to explain a type of effect, one feels a sense of thorough dissatisfaction. Even when a very precise symmetry seems to require the postulation of a specially instituted and artificially devised arrangement, one is still inclined to regard it as the necessary result of more general laws and to continue to observe the rule of unity, before resorting to an explanation in terms of an artificial provision. (AA II: 113)

For many kinds of natural formations, Kant is willing to allow for their dependence on a limited set of natural causes and principles alone. He realizes, however, that the matter is somewhat more complicated in the case of organic form:

[I]t would be absurd to regard the initial generation of a plant or animal as a mechanical effect incidentally arising from the universal laws of nature; nonetheless, there is a two-fold question, which has remained unanswered for the reason mentioned. Is each individual member of the plant- and animal-kingdoms directly formed by God, and thus of supernatural origin, with only propagation, that is to say, only the periodic transmission for the purposes of development, being entrusted to a natural law? Or do some individual members of the plant- and animal-kingdoms, although immediately formed by God and thus of divine origin, possess the capacity, which we cannot understand, actually to generate their own kind in accordance with a regular law of nature, and not merely to

unfold them? There are difficulties on both sides, and it is perhaps impossible to make out which difficulty is the greatest. (AA II: 114)

The dilemma Kant construes here is: do we violate the order of nature by allowing a variety of phenomena to be unexplainable by the principles of nature or do we violate it by postulating new principles of nature? In other words: do we add the surplus of information for which we cannot account by postulating it in the initial state of the thing (its origin), or do we add it by allowing for new transformative capacities in nature (the process of development)? This dilemma is forced upon us by the specific mechanical unintelligibility of organic generation:

in the light of everything we know, it is utterly unintelligible to us that a tree should be able, in virtue of an internal mechanical constitution, to form and process its sap in such a way that there should arise in the bud or the seed something containing a tree like itself in miniature, or something from which such a tree could develop. The internal forms proposed by *Buffon*, and the elements of organic matter which, in the opinion of *Maupertuis*, join together as their memories dictate and in accordance with the laws of desire and aversion, are either as incomprehensible as the thing itself, or they are entirely arbitrary inventions. (AA II: 115)

Kant here criticizes Buffon and Maupertuis for failing to offer a really mechanical theory of generation. As I explained in section 2.2.3 of this chapter, Buffon and at least the early Maupertuis intended to offer a mechanist theory by expanding the all-too-restricted set of mechanical principles employed by the followers of Descartes. Kant seems to be saying, however, that the analogy with Newton fails. Indeed, mechanists were reticent towards allowing new forces and principles in their austere ontology and explanatory apparatus. They could, however, be forced to admit gravity because gravity is utterly general and reinforces mechanical explanations. The same does not seem to hold for Buffon's and Maupertuis's principles, which are, in Kant's appraisal, incomprehensible and/or arbitrary. The internal forms would be incomprehensible because they refer to a kind of non-extended structuring, and whilst gravity can be regarded as a penetrating force, working on the interior of matter, it is exhausted by its extensive effects (motions, accelerations, etc.). However, even if one were to allow for such interior properties, they would still be arbitrary, because unlike gravity, they do not operate generally, but are specific to species or individuals, and are selective in their operations. This selectivity in effect, however, is precisely what is in need of explanation.

But this does not mean that Kant therefore embraces preformation:

In this case, the origin of all such organic products is regarded as completely supernatural; it is, nonetheless, supposed that the natural philosophers have been left with something when they are permitted to toy with the problem of the

manner of gradual propagation. But consider: the supernatural is not thereby diminished, for whether the supernatural generation occurs at the moment of creation, or whether it takes place gradually, at different times, the degree of the supernatural is no greater in the second case than it is in the first. The only difference between them relates not to the degree of the immediate divine action but merely to the *when*. As for the natural order of unfolding mentioned above: it is not a rule of the fruitfulness of nature, but a futile method of evading the issue. For not the least degree of the immediate divine action is thereby spared. (AA II: 115)

I believe this passage is best interpreted as a criticism of the arguments in favor of preformation that led Justin Smith to postulate a parallel between preformation and pre-established harmony (cf. 2.2.2). Remember that authors such as Leeuwenhoek argued that, if the mechanical generation of living systems is unintelligible, this means that we must admit that all organisms are created directly by God. The theory of preformation is more attractive because it does not require constant miraculous intervention by God in order to explain how organisms occur in the world. It just assumes that all organisms were created at the moment when God created the universe. Kant was dissatisfied with this argument because he regarded it as a mere sleight of hand. To see why this is so, we need to elaborate the parallel between occasionalism and pre-established harmony on the one hand and continuous creation and preformation on the other.

Leibniz often stresses the superiority of pre-established harmony over occasionalism. The reasoning behind this assumption of superiority is unclear, however, for a reason that was noted by Antoine Arnauld in his correspondence between Leibniz:

With regard to that which you say, namely that even though my arm lifts when I want to lift it, this is not because my soul causes this movement in my arm; but that this is because when I want to lift it, this is precisely at the moment at which everything in the body is disposed for this effect, such that the body moves in virtue of its proper laws, even though it comes to pass through the admirable but unwavering agreement of things amongst each other, that this law conspire to this effect at exactly the moment when the will is brought to it. God having taken it into account in advance, when he resolved to this sequence of all things in the universe. It seems to me that this is saying the same thing in other words as is said by those who pretend that my will is the occasional cause of the movement of my arm, and that God is its occasional cause. For they do not pretend that God does that *within time* by a new volition, which he has *every time* I want to move my arm; but by this sole act of eternal volition, through which he has willed to do everything that he has foreseen as necessary in order that the universe be thus as he has judged that it had to be. (G II :84; my translation and stress)

Arnauld protests that Leibniz has no right to claim the superiority of pre-established harmony over occasionalism because it requires less intervention in the order of nature on God's part, since the occasionalists agree with Leibniz that the occasioning of causes is always due to general volitions, and not particular interventions. As the terms in italics indicate, Arnauld denies that occasionalism is committed to the fact that God intervenes *within time*, that all his interventions are contemporary, i.e. at the moment of the miracle of creation. In his response to this criticism, Leibniz concedes that occasionalism is not inferior because it requires more volitions and creations on God's part, but because it requires that all powers reside in God:

If I properly understand the views of the authors of the occasional causes, they introduce a miracle which is no less miraculous for being continual. For it seems to me that the notion of miracle does not consist in rarity. One might say in this matter God acts only according to a general rule, and consequently he acts without miracle. But I do not grant that consequence, and I believe that God can make general rules for himself even with respect to miracles. For example, if God had resolved to give his grace immediately or perform some other action of this nature every time a certain condition was satisfied, this action, though ordinary, would nevertheless still be a miracle. I admit that the authors of occasional causes might give another definition of the term, but, according to common usage, it seems that a miracle differs internally and substantively from the performance of an ordinary action, and not by the external accident of frequent repetition; properly speaking, God performs a miracle when he does something that surpasses the forces he has given to creatures and conserves in them. (G II: 92-93; Leibniz 1989: 82-83)

This passage reveals that the distinction Leibniz sees between pre-established harmony and occasionalism is not that the latter is less economical than the former, but that the former preserves a sense in which nature itself is dynamic. Leibniz advances, against Malebranche and other occasionalists, that the superior position is still the Thomist orthodoxy that God creates and sustains creatures with their causal powers, even though he claims that these causal powers only truly operate within these creatures. Occasionalism robs nature of this autonomous power by stating that creatures do not have causal powers, and that all causal powers reside in God. For Leibniz, then, creatures have causal powers, and these are not exerted on each other, but do harmonize with each other, whereas for occasionalists (according to Leibniz), creatures lack causal powers and the harmony between causes and effects is due to God's general volitions.

This gives us an important clue as to what Kant is trying to argue in the above-quoted passage. He is, I believe, arguing that preformation is better than continuous creation only on the equivocation of the concept of a miracle noted by Leibniz. After all, preformationists are ultimately equally committed to the fact that the creation of a

living system requires a distinct act of creation by God. They only disagree over the fact whether this distinct act of creation takes place in time, or at the moment of creation. And at this point the analogies between occasionalism and continuous creation and pre-established harmony and preformation break down. On the occasionalists notion of a miracle, the degree of supernaturalism could be decreased only if not every single generation required a separate creation, but could instead take place through a general principle. But this is of course exactly what preformation denies: no generation can take place according to a general principle. On the Leibnizian version of a miracle, on the other hand, the parallel comes out little better. Leibniz would agree that every organism is individually preformed by God, but argue that the subsequent development of the organism could be due to proper powers of that organism. This, however, arguably makes *development* less of a miracle, but not *generation*. In order for generation to cease to be a miracle on either of these accounts, it would require that God endow upon his creatures their own proper powers to produce their like, and thereby establish through a general volition the possibility of all future generations. Again, this is precisely what Leibniz wants to deny.

That this analysis is what Kant had in mind is supported by the fact that he concludes the passage with a description of the position, rejected by all supernaturalists, that I just described: “[t]he purpose of these considerations has simply been to show that one must concede to the things of nature a possibility, greater than that which is commonly conceded, of producing their effects in accordance with universal laws.” (AA II: 115). This position is less supernaturalist on both notions of a miracle, because it allows for a general volition for each species, and for every individual of this species to truly have the natural power to propagate. Moreover, this position gives rise, further on in the *Only Possible Argument*, to a rule of Kant’s revised version of physico-theology, which stresses both the unity and the causal autonomy of nature:

One will presume that the necessary unity to be found in nature is greater than strikes the eye. And that presumption will be made not only in the case of inorganic nature but also in the case of organic nature as well. For even in the case of the structure of an animal, it can be assumed that there is a single disposition [*Anlage*], which has the fruitful adaptedness to produce many different advantageous consequences. Initially, we may have supposed that a variety of special provisions must have been necessary to produce such effects. Careful attention to the necessary unity of nature is both consonant with philosophy and advantageous to the physic-theological method of inference. (AA II: 126)

By themselves, none of these passages point towards the concrete nature of Kant’s own theory of generation. In fact, the only positive comments he makes are restrictions on the kind of theory that he would want to accept. Nevertheless, Kant reveals himself to be clearly dissatisfied with both metamorphosis and preformation because of their

inability to account for natural change and generation without stretching the meaning of “natural” beyond the bounds of plausibility. He also introduces a core term of his own approach: that of disposition or *Anlage*. Here, it is important to indicate that he speaks *only* of disposition, and that the structure of animal can be accounted for by a *single* disposition, since the evidence for Kant’s preformationism rests on the fact that this term is to be understood in tandem with that of germ (*Keim*) (Cf. Sloan 2002 239-240). It is telling, in this respect, that the term first appears autonomously in the early 1760s, and that the talk of “germs” is a later development, one from the 1770s. In the next section, I will argue that, although it is true that the term germ is to be understood in its systematic connection with that of a disposition, it is the latter term, and its origin in an explicitly anti-preformationist theory of generation, that maintains the upper hand.

### 2.3.2 Races and Dispositions

Between the Inaugural Dissertation of 1770 and the Critique of Pure Reason of 1781 lie what we are now used to calling Kant’s silent decade. This term is somewhat of a misnomer, since Kant did publish some papers in this period, although none of them have even remotely attained the status of the works of the 1780s and the 1790s, the so-called critical period. These texts offer several interpretative difficulties, not in the least because it is unclear whether they are to be understood as pre-critical or as already critical. We may be tempted to regard them as pre-critical because they continue a project in which Kant was engaged since the late 1750s, that of physical geography and anthropology. On the other hand, their themes persist into works from the critical period. This gives us the additional task of ascertaining whether the ideas expressed in them underwent any changes due to the critical turn.

Phillip Sloan and John Zammito have argued that the works of the 1770s are continuous not with the critical writings as a whole, but with the early critical period. They regard them as expressive of Kant’s endorsement of preformationism, which allegedly continued until the mid 1780s, when, under the influence of his confrontations with Johann Gottfried Herder’s and Johann Friedrich Blumenbach’s writings, he was forced to develop the theory now known as generic preformationism, which is a more mitigated form of preformation than the earlier one. As we will see, this assumption is certainly not outrageous, but it does face serious difficulties.

The most important element of Kant’s turn to preformation in the 1770s is, according to Sloan and Zammito, his usage of the term Germ (*Keim*) along that of Disposition (*Anlage*). The former term appears explicitly for the first time in the 1771 review of Moscati’s work *Of the Corporeal Essential Differences between the Structure of Humans and Animals*. Kant’s review opened with the following sentences:



Here we have again the natural human being on all fours, to which he is returned by an astute anatomist, which *Rousseau* as a philosopher did not succeed in doing. Dr. *Moscatti* proves that the upright gait of the human being is contrived and against nature; that he is indeed built to maintain himself and move about in this position; but that, if he makes this his necessity and constant habit, discomforts and maladies result which demonstrate sufficiently that he was enticed by reason and imitation to deviate from the first, animal set-up. (AA II: 423)

This passage already reveals the specific interest with which Kant read this work: that of the origin of man.<sup>7</sup> The reference to *Rousseau* shows the concern with this issue and the extent to which rationality is natural or unnatural to man. The term germ pops up in the closing passages of the review as a response to precisely this question:

The first foresight of nature was that the human being as an animal be preserved *for himself and his kind*; and for that the position which is most suited to his internal build, the situation of the fetus and the preservation in dangers is the *four-footed* one; but that there also has been placed in him a germ of reason through which, if the latter develops, he is destined *for society*, and by means of which he assumes permanently the most suitable position for society, viz., the *two-footed* one. Thereby he gains, on one side, infinitely much over the animals, but he has also to live with the discomforts which result for him from the fact that he has raised his head so proudly above his old comrades. (AA II: 425)

This passage seems to ascribe to man a natural origin as an animal amongst animals, and seems to speak of this origin as if it had occurred at some point in history. Yet, Kant here also ascribes to man a germ of reason, through which he is destined for unnatural life, for a life that is at odds with the natural way of living. In the next chapter, it will become clear that we are dealing here with the first formulation of Kant's distinct philosophy of history as it will be developed in the critical period.

More important for our present purposes, however, is the paper *Of the different Races of Human Beings*, published first as an announcement of his lectures on physical geography in 1775, and again in a revised version as an article in 1777. In this paper, the distinction between germs and dispositions appears explicitly, although here again it is useful to address the specific content of the paper in more detail in order to avoid misunderstandings.

The opening passages of the paper are clearly reminiscent of several themes from the *Only Possible Argument*:

---

<sup>7</sup> In this quote, we can already see the tendency towards a recapitulation theory, which maintains that the same sequence governs the development of the species and that of the individual. Here, Kant does not address this, but it will work on in his philosophy of history (cf. chapter 3).

The natural division into species and kinds in the animal kingdom is grounded on the common law of propagation, and the unity of the species is nothing other than the unity of the generative power [*zeugenden Kraft*] that is universally valid for a certain manifoldness of animals. For this reason, Buffon's rule, that animals which produce fertile young with one another (whatever difference in shape there may be) still belong to one and the same physical species, must properly be regarded only as the definition of a natural species of animals in general in contrast to all school species of the latter. The school division concerns *classes*, which divide the animals according to *resemblances*, the natural division concerns *phyla*, which divide the animals according to *relationships* in terms of generation. The former provides a school system for memory; the latter provides a natural system for the understanding. The first only aims at bringing creatures under titles; the second aims at bringing them under laws.

According to this concept, all human beings on the wide earth belong to one and the same natural species because they consistently beget fertile children with one another, no matter what great differences may otherwise be encountered in their shape. One can adduce only a single natural cause for this unity of the natural species, which unity is tantamount to the unity of the generative power that they have in common: namely, that they all belong to a single phylum, from which, notwithstanding, their differences, they originated, or at least could have originated. In the first case, human beings belong not merely to one and the same *species*, but also to one *family*; in the second case they are similar to one another but not related, and many local creations would have to be assumed – an opinion which needlessly multiplies the number of causes. (AA II: 329-330)

Here we find the same call for unity in explanation in nature as in the *Only Possible Argument*: Kant argues that we must understand all human races as being of the same natural species, because it is undesirable to postulate specific further causes merely in order to account for the minor differences between them. Kant endorses, then, a monogeneticism based on the genealogy of man that could be regarded as a standard creationist story, much as Leibniz did before him (Cf. Smith 2011: 269-274). His way of dealing with this monogeneticism, however, cannot be the same as that of Leibniz, because of the rationale lying behind his arguments.

As we saw, in the *Only Possible Argument*, Kant protested against preformation because it had to postulate a distinct creation, a distinct cause or principle, for each biological individual. Given Kant's unwillingness to admit multiple causes to account for multiple human races, I think it is fair to assume he would still have been reticent towards the idea of individual preformation. This is further attested to by the main focus of this passage: *the unity of the generative force*. Recall that already in the 1763 writing, he indicated that "one must concede to the things of nature a possibility, greater than that which is commonly conceded, of producing their effects in accordance with universal laws" (AA II: 115) and that in the case of the structure of living beings,

“there is a single disposition [*Anlage*], which has the fruitful adaptedness to produce many different advantageous consequences” (AA II: 126). What he seems to be concerned with here is to show that “some individual members of the plant- and animal-kingdoms, although immediately formed by God and thus of divine origin, possess the capacity, which we cannot understand, actually to generate their own kind in accordance with a regular law of nature, and not merely to unfold them” (AA II: 114).

In contradistinction to preformationism, this account does not imply that individuals are preformed, but rather that species have the capacity to regularly produce offspring sharing their specific marks and their specific generative force. But this leaves Kant with the task of explaining from where the specific differences and the specific range of differences within species come. It is only for this task that he introduces the notions of germs and dispositions:

The grounds of a determinate unfolding which are lying in the nature of an organic body (plant or animal) are called *germs*, if this unfolding concerns particular parts; if however, it concerns only the size or the relation of the parts to one another, then I call them *natural predispositions*. [...] In the migration and transplanting of animals and plants it creates the semblance of new kinds; yet they are nothing other than variations and races of the same species the germs and natural predispositions of which have merely developed on occasion in various ways over long periods of time.

Chance or the universal mechanical laws could not produce such agreement. Therefore we must consider such occasional unfoldings as *preformed*. Yet even where nothing purposive shows itself, the mere faculty to propagate its adopted character is already proof enough that a particular germ or natural predisposition for it was to be found in the organic creature. For outer things can well be occasioning causes but not producing ones of what is inherited necessarily and regenerates. As little as chance or physical-mechanical causes can produce an organic body, just as little as chance or physical-mechanical causes can produce an organic body, just as little will they add something to its generative power, i.e., bring about something that propagates itself, if it concerns a special shape or relation of the parts. (AA II: 434-435)

This passage is very important, even if it were only for the misconceptions that it may bring about. Sloan (2002: 233-235) goes through great pains to show that the term “germs” stems unambiguously from classical preformationism, which would explain why Kant believed that “we must consider such occasional unfoldings as *preformed*”. Zammito (2003: 83) has concluded from this analysis that “[t]he specific form of *preformation* that Kant endorsed was the sophisticated version developed by Bonnet and Haller in the early 1760s in response to the challenge first of Maupertuis and Buffon and then, more fundamentally, of Caspar Friedrich Wolff”. I believe this is mistaken, since it cannot be made to square with the text.

As argued in section 2.2.3., the preformationism of Haller and Bonnet was indeed sophisticated, but no less a preformation. All preformationists believed that environmental and sometimes even hereditary effects could influence the outcome of the process of unfolding. But they also believed that these effects were extraneous to the individual germ itself, which had been preformed at creation. Secondly, they obviously made no distinction between the germ and something like a disposition. In fact, they believed there to be a single germ which was then unfolded on occasion of incubation and appropriate conception and nutrition. This is certainly not what Kant is talking about here.

Kant's concern in this passage is more with the unity and constancy of the generative force. He suggests that this unity and constancy is required in order to account for the reproductive and restorative (and probably nutritive) properties of living beings. But he equally wants to account for the adaptive nature of this property, and the inheritability of this adaptedness. As a result, he postulates that there are two kinds of internal constraints on the generative force that determine the outcome of the developmental process. The first kind he calls germs, and he considers to be the factors determining the development of the animal through the generative force. The second kind are not as determinate, but instead determine which possible changes can occur to the constraints of the generative force. The most original aspect of his position is that he believes that not just the stabilities, but also the changes in form, are due to internal constraints rather than due to outside occasions. This is a major shift away from preformation, which suggests that the preformed structure is fully determined, and all real variations are due to outside influences. What is "preformed", according to Kant, is not a fixed structure, but a constraint on the forces and processes of generation and development. We will see in the course of this chapter and in the course of this dissertation which kinds of constraints Kant has in mind.

This is related to another major aspect of the position Kant advances in *Of the different Races of Human Being*, namely its departure from classical systematics. As Kant notes in the opening paragraph of the essay:

The school division concerns *classes*, which divide the animals according to *resemblances*, the natural division concerns *phyla*, which divide the animals according to *relationships* in terms of generation. The former provides a school system for memory; the latter provides a natural system for the understanding. The first only aims at bringing creatures under titles; the second aims at bringing them under laws. (AA II: 329)

Kant announces here that he intends the account of species in terms of a generative force to be a departure from the traditional account. I believe this passage continues the criticism of preformationism offered in the *Only Possible Argument*. To see why this is so, we need to briefly return to the preformationist account of genealogy. It is clear that

preformationism cannot maintain that one creature is a parent of the other if the former has produced the latter through its natural capacities, since it denies that creatures are naturally produced, and *can* be naturally produced. This brings up the following question: in what sense can we then say that one creature engendered another, if both creatures are strictly speaking co-originary? I know of two preformationist answers to this question. The first answer is that of the theory Charles Bonnet called *emboîtement*, namely that God placed, at the moment of creation, all children in the reproductive organs of their parents, to develop there on the occasion of fertilization. The reason for placing a creature in another creature is that they resemble one another at least to the extent that they are of the same species. The other answer is due to the theory Bonnet called *dissémination*. According to this theory, the preformed organisms are dispersed throughout nature, and develop only when 1) they happen to enter a creature that sufficiently resembles it and 2) the latter creature is fertilized by a creature that sufficiently resembles it.

The upshot of both of these answers is that we are not of the same species as our parents because we were engendered by them, but that we are engendered by our parents because we are of the same species as them. The notion of a species is therefore exhausted by a degree of resemblance between otherwise independently generated, co-originary entities. Any further resemblance, such as the typical constancy of traits in families that we now tend to regard as hereditary, are purely external influences on the embryo due to the specific physiological (most importantly nutritional) context in which it happens to develop (i.e. the body of the mother).

It is now easy to see how Kant's account is a criticism of preformationism. For Kant, preformation can never found a species concept based on a natural kind, since it cannot find a principle for the similarities between members of a different species. In absence of such a principle, we can never know which degree of similarity is required to speak of the same species, and we can never know which similarities are required for a specific identity. In absence of this, we will always be forced to classify natural entities on the basis of their surface similarities, realizing that these surface similarities can always be challenged by another choice of criterion.

Kant would therefore regard his own account as superior because it postulates a general principle responsible for the specific identity, namely the formative force. This formative force is *specified* and *specifies* by means of the germs and dispositions. On the basis of the assumption of such a species identity, we can identify which degree of similarity and which similarities are required to speak of an identity between species. We can also reject the idea that similarities are required to speak of species identity; all that is required is the continuity of production, and no specified amount of dissimilarity can lead one to reject the idea that the same productive force was at work. Finally, as we already saw, Kant's account entails that the changes in the productive force are internal

changes, not external accidents. The productive force is therefore characteristically plastic.<sup>8</sup>

## 2.4 Epigenesis in the Critical Period

### 2.4.1 Germs and Dispositions as Internal Limitations

John Zammito is famed for addressing the great importance of Kant's confrontation with Johann Gottfried Herder's *Ideen zur Philosophie der Geschichte der Menschheit*, the first volume of which appeared in 1784. According to Zammito (1992: 181) "Herder's grand project in the *Ideen* was to find how man as a creature of nature figured in man as an artifice of culture, to read these two dimensions of man as in continuity". In his efforts to situate man within the totality of nature, Herder interprets organic nature as a continuity of organizational forms, where (anticipating Geoffroy Saint-Hilaire and Cuvier) the various morphologies are considered as variations on a single basic prototype (Herder 1869 : 49). From this he considers himself entitled to conclude "that Nature, in the infinite variety that she loves, seems to have formed all life on our earth according to one main plasma" (Herder 1869: 49; my translation). A bit further in the work, Herder boldly states that this process of formation (*Bildung*) is epigenesist rather than a mere unfolding of preformed germs (*Keime*):

The theory of germs, that has been adopted in order to explain vegetation, doesn't really explain anything at all; for the germ is already in the product, and where that is, there must already be an organic force that produces it. No analyst would have discovered all future germs in the first grain of creation; they do not become visible until the plant has reached its full force; and from all our experiences we do not have the right to attribute it to anything else than an organic force of the plant itself, working on it with quiet intensity. (Herder 1869: 63; my translation)

Herder's version of epigenesis stresses the organizational power over the actual structure, the process of formation (*Bildung*) over the form itself. This power is identical behind the many forms that it produces and constitutes the single organizational nexus that suffices to engender and understand the whole of nature: it is nature considered as active, as *natura naturans* rather than *natura naturata*.

---

<sup>8</sup> We will return to the topic of plasticity in 4.2.2

Zammito has made much of Kant's resistance towards these ideas, and interpreted it as an indication of Kant's refusal of epigenesis altogether. But this is not so clear from Kant's actual objections in his 1785 review of the first two parts of Herder's work:

[Herder] calls the cause of the climatic difference of human beings a *genetic* force. The reviewer has the following concept of the meaning of this expression, in the author's mind. He wants to dismiss on the one side the system of evolution and yet also on the other side the merely mechanical influences of external causes as providing unworkable grounds of elucidation, and he assumes as its cause a principle of life, which appropriately modifies *itself* internally in accordance with differences of the external circumstances; with this the reviewer fully concurs, only with this reservation, that if the cause organizing itself *from within* were limited by its nature only perhaps to a certain number and degree of differences in the formation of a creature (so that after the institution of which it were not further free to form yet another type under altered circumstances), then one could call this natural vocation of the forming nature also "germs" [*Keime*] or "original predispositions" [*ursprüngliche Anlage*], without thereby regarding the former as primordially implanted (as in the system of evolution), but merely as limitations, not further explicable, of a self-forming faculty, which latter we can just as little explain or make comprehensible. (AA VIII: 62-63).

Mind that Kant's hesitance is directed not to the idea of a genetic force, which is similar to Kant's own generative force, nor to the fact that this generative force is plastic and can modify itself internally. Like Kant, Herder refuses the preformationist system and the system of metamorphosis. The distinction only comes in once this genetic force needs to be qualified. Herder attributes to it an almost unlimited plasticity, which allows it to give rise to all kinds of structures over time. It is this latter assumption with which Kant disagrees: according to him, the generative force must be internally constrained in order to be explanatory at all. It is in this context that he revisits the preformationist-sounding terminology of the 1770s and the early 1780s, interpreting them "merely as limitations, not further explicable, of a self-forming faculty, which latter we can just as little explain or make comprehensible". I submit here that this view of germs may very well have been the one he held all along: in all previous passages, the germs and predispositions failed to match up with classical preformationist notions, and appeared only as qualifications and internal limits on Kant's central concept: that of the unitary generative force.<sup>9</sup> The question remains why Kant then continued to use this deceptive

---

<sup>9</sup> Peter McLaughlin (2007: 287) commented on Kant's use of the notions of germs and predispositions as follows: "[Kant's] germs and predispositions do still fit the determinist eighteenth-century model of pre-given potentialities". This is true, but it does not seem to be the case that they imply anything more than the

terminology.<sup>10</sup> This is a question that I will not be able to answer until the end of the next chapter.

## 2.4.2 Generic Preformationism and Transcendental Epigenesis

The final text to which I would like to draw attention in this chapter is the one in which Kant coins his term “generic preformationism” and in which he approvingly refers to Blumenbach’s recently advanced concept of Formative Drive (*Bildungstrieb*). This text is §81 of the *Critique of the Power of Judgment*, where Kant tackles the teleological principle as it is applied to the case of generation. He discerns two distinct ways of accounting for this principle:

If the teleological principle of the generation of these beings (i.e. natural purposes) is assumed (as cannot but be the case), then the cause of their internally purposive form can be grounded in either *occasionalism* or *pretabilism*. According to the former, the supreme world-cause, in accordance with its idea, would immediately provide the organic formation to the matter commingling in every impregnation; according to the latter, it would only have placed in the initial products of its wisdom the predisposition by means of which an organic being produces more of its kind and constantly preserves the species itself, in which a nature that continuously works at their destruction simultaneously makes good the loss of the individuals. (AA V: 422)

Kant here assumes that the specificity of organic form cannot be produced by the most general of mechanical principles alone, a thesis for which he argued in the preceding paragraphs of his *Critique of Teleological Judgment*. Additionally, he assumes that hylozoism or vital materialism is not an option. These two assumptions echo those of the 1763 discussion, where Kant rejects not only classical mechanist accounts of generation, but also the attempts of theorists like Buffon and Maupertuis to account for complexity by making (in Kant’s opinion) ad hoc additions to the principles and causes at work in nature. For this reason, he is bound to accept either the occasionalist version or the pretabilist one. Kant understandably immediately rejects the occasionalist account (AA V: 422), because it would make generation fully miraculous and unnatural. This leaves him with what he calls pretabilism:

---

eighteenth-century model of *determining* factors, and there is surely a great difference between a preformed element and a determining factor.

<sup>10</sup> Actually, it is not as deceptive as we would initially think. Johannes Nikolaus Tetens had already written in 1777 that the term “germ” or “keim” is neutral with respect to preformationism or epigenesis, since both theories have to admit of some organization (Tetens 1777, II: 455; cf. also further on in 6.2.3).



Now *pretabilism* can in turn proceed in two ways. Namely, it considers each organic being generated from its own kind as either the *educt* or the *product* of the latter. The system of generatings as mere educts is called that of *individual preformation* or the *theory of evolution*; the system of generatings is called the system of *epigenesis*. The latter can also be called the system of *generic preformation*, since the productive capacity of the progenitor is still preformed in accordance with the internally purposive predispositions that were imparted to its stock, and thus the specific form was preformed *virtualiter*. Given this, the opposing theory of individual preformation might better be called the *theory of involution* (or that of encapsulation). (AA V: 422-423)

At last we see Kant explaining the remnants of preformationist language in his own theory of epigenesis. He has been assuming that the productive capacity that accounts for epigenetic production has a specific structure, and that this structure is to be understood as expressing a virtual preformation rather than an actual one. The process of epigenesis is then one that articulates and differentiates organic form that is in itself homogeneous, but that must be considered as virtually containing the articulations and differentiations required to yield its full-formed structure. Here, I would like to remind the reader that the assumption of virtual rather than actual initial differentiation was exactly the position that both preformationists and metamorphosists rejected, and that epigenesists endorsed. For neither Harvey nor Wolff, epigenesis meant an unrestricted or blind force: they believed instead that the required restrictions were not pre-existing structures, but rather proper to the process and the circular causality (cf. 2.2.1) of generation. Here, Kant coins the term “virtual” in order to assign a status to the specific way in which the guiding idea of the whole pre-exists the parts of that whole.

It is within the context of such a position that Kant offers the following praise of Blumenbach:

He [Blumenbach] begins all physical explanation of these formations with organized matter. For he rightly declares it to be contrary to reason that raw matter should originally have formed itself in accordance with mechanical laws, that life should have arisen from the nature of the lifeless, and that matter should have been able to assemble itself into the form of a self-preserving purposiveness by itself; at the same time, however, he leaves natural mechanism an indeterminable but at the same time also unmistakable role under this inscrutable *principle* of an original *organization*, on account of which he calls the faculty in the matter in an organized body (in distinction from the merely mechanical *formative power* [Bildungskraft] that is present in all matter) a *formative drive* [Bildungstrieb] (standing, as it were, under the guidance and direction of that former principle). (AA V: 424)

Kant’s respect for Blumenbach’s position centers on several interesting features of this account. For one, Blumenbach always starts from organized matter. This means that he

takes the original matter of the living being to be not a mere chaotic mixture of elements, nor a homogeneous material, but to be in some way already organized. This organization is not characterized in terms of its resemblance to the later stage of that organization, and therefore not as its preformed version, but rather as the organization giving rise to the latter through a genuinely productive force: the *Bildungstrieb*. The benefit of such an analysis is that it allows us to admit that order does not arise out of chaos, whilst still maintaining that the order that arises is not fully predetermined by the order out of which it arises.

In addressing the question whether this position of Kant's is preformationist or epigenesist, we need to note that it may not be far removed from Aristotle's position after all, as is revealed by the following comment by Balme:

Aristotle, like Plato and probably all the ancients, sees the primary actions of the elements as quantitatively indeterminate until limit is imposed upon them. If limit is naturally imposed, not at random but regularly and usually, it must come from a pre-existing source; this he interprets as an enclosed system of movements in air earth fire and water; and the only place where it can be is in the parent. [...] It is summed up in the cardinal principle that 'a man begets a man'. (Balme 1987: 283)

Balme describes Aristotle's theory as implying that the development of the fetus through an ordered process is due to the fact that there are limits or constraints placed upon the material processes that make up the process of generation. These limits or constraints are themselves due to the pre-existence of an organic structure with similar features (i.e. of the same kind). It would be disingenuous to take the word "pre-existence" here in a way that suggests any of the pre-existence of preformation theories of Early Modernity. Rather, it means that the specific (i.e. species-) constraints are operative because reproduction is always to be understood as a power of a living system, the parent.

It is not difficult to see the parallels between this view and Kant's. Kant too emphasizes that what we should regard as pre-existing is not the individual, but the species. This can easily be understood as the claim that in order for a living system to be generated, it is required that there already exist another living system which embodies the specific constraints of that species. Moreover, as I have argued, the notions of germs and predispositions are best understood as constraints on the developmental process, and the developmental process as the exercise of the formative force. The formative force, then, acquires its specificity because it is subject to, and incorporates, the specific constraints exemplified by the generating structure (the parent).

This brings us to the real meaning of Kant's decried species fixism. The point behind this is not that God has ordained and maintained the number of species. It is rather that removing the stability of species-characteristics in development would make development unintelligible. In stressing the plasticity and changeability of living form,

we should not lose out of sight the idea of specification and speciation, which require that we explain this change by invoking the relations between specific structures. A general, unrestricted generative force like that hypothesized by Herder is not capable to give rise to all the species – because it fails to account for the specification and the specificity, it is in fact incapable of giving rise to *any* species. Additionally, Kant is not optimistic about origin of life theorizing, because he is committed to the fact that every living being requires a pre-existing living being. Hence, life requires either an eternal regress or a moment of brute creation. But neither of these options need to be endorsed by the father of transcendental idealism, since he has shown both alternatives to be avoidable.

## 2.5 Conclusion

In this chapter, I have argued that Kant maintained, throughout the whole of his career, a characteristically epigenesist theory of generation, and have argued against the idea that Kant was ever tempted by preformationism. I have done this by showing that much of the considerations that have been taken as concessions to preformationism are best understood as attempts to refine his own picture of epigenesis, and do not constitute major divergences from the epigenesist tradition. This is important because Kant also employed the language of his epigenesist theory of development in his philosophy of history and in his epistemology. I will discuss the meaning of this employment of embryological metaphors in the following chapter. In order to set the stage for my discussion there, however, I believe it is appropriate to conclude here by indicating some of the crucial characteristics of Kantian epigenesis that are invoked in his philosophy of history and epistemology.

First of all, Kant's version of epigenesis is meant to account for the continuity of species. Along with Buffon, Kant insisted that species are determined by interfertility, and by heredity. Two living beings are of the same species when they somehow share a genealogy. By invoking this idea, Kant is stressing the continuity of living form, namely the idea that living beings are always brought forth by other living beings from the same kind.

Second, Kant stresses that this continuity expresses itself in the operation of the living force and the process of development, not in the internal connection with the previous structure. He is dissatisfied by preformationism because it assumes that the structure required for explaining a later structure need be both numerically and structurally identical with it and similar to it. For Kant, all that is required is that a previous structure constrains the production of the latter structure, and that these

constraints are virtually present in the productive force rather than actually present in the producing structure.

Third, Kant's theory of predispositions is not only meant to account for the continuities, but also for the changes. The concept of a predisposition serves to do this in two ways. On the one hand, it stresses that, although an earlier structure is for a second structure of the same kind to arise, the similarity between the two structures can be rather low. All that need to be shared are the basic germs and predispositions. For Kant, then, two structures relate not as an enveloped preformed structure and a developed formed structure would, but as parent and child – sure, they share species traits, but the individuality is not the same, and hence not everything is determined. The capacity of parents to bring forth children is not, for him, the capacity to occasion preexisting individuals – it is the capacity to bring forth new individuals, with their own individual traits. This capacity is due to a fundamentally plastic generative force.

Finally, the idea of germs and predispositions makes room for the contingent without thereby having to externalize history. In his writings on races, Kant was concerned with showing that environmental influences could only effect changes in formative forces if they somehow occasioned the predispositions lying within those forces. This means that he did believe environmental influences could lead to changes, but that they could not do so through direct influence.

All of these characteristics of the theory are essential to a proper understanding of the way in which Kant's philosophy can integrate the historical, both in his philosophical anthropology and his epistemology. To see what they mean in these contexts, we must now turn to the discussion of the theory of epigenesis as it appears in the writings that are not (expressly) concerned with the life sciences.

## Chapter 3

# Epigenesis and History

*Ursprung ist das Ziel*  
- Karl Kraus

In this chapter, I argue that Kant's usage of language from embryology in other doctrines of his is best understood on the basis of the version of epigenesis ascribed to him in chapter 2. By reading his philosophy of history and his epistemology through this epigenesist lense, we see that Kant was indeed open to the historical and the dynamic. In section 3.1, I sketch the now dominant view of the implications of Kant's embryology for his philosophy of history and his epistemology, which denies this historical and dynamic element. In section 3.2, I argue against this view in the case of Kant's philosophy of history. I start off, in 3.2.1, by sketching what I take to be the target of Kant's views: Rousseau's theory of human nature. In the next two subsections, I argue that Kant's epigenesis was meant to undermine the Rousseauist picture of man. In 3.2.2, I argue that the major difference between Kant and Rousseau is that the former believed the essence of man to be plastic, whereas the latter insisted that it is fixed. In section 3.2.3, I argue that the germs and dispositions Kant attributes to man are best understood as those capacities which make us capable of history and of change, rather than those which for ever determine our limits and essences. In section 3.3, I argue that Kant meant his epigenesis in epistemology as well. In 3.3.1, I argue against the thesis that the embryological language in the first edition of the *Critique of Pure Reason* is still preformationist, and can instead be read as in harmony with the view ascribed to Kant in chapter 2. In 3.3.2, I then argue that the embryological language of the second edition is clearly epigenesist, rather than mildly preformationist. Together, these arguments show that there are good reasons for ascribing to Kant a more dynamic view of history and epistemology. In section 3.4, I make a first attempt at probing what such a view might entail.

### 3.1 Epigenesis Beyond Generation

Kant's theory of epigenesis is of central importance because he invokes it in the context of his writings on philosophy of history, philosophical anthropology and pedagogy, and in the *Critique of Pure Reason*. On my account, this use of embryological metaphors is essential because it refers to the capacity of Kant's philosophy to deal with the historical and with historicity. Ironically, then, this very link is now usually denied because of the wide-spread view, best expressed in the accounts given by Sloan and Zammito, that Kant's theory of generation was actually more preformationist in nature, and hence like preformationism fundamentally ahistorical. In this section, I will indicate the implications of this for the interpretation of Kant's philosophy of history and his epistemology. I will then go on to offer an alternative interpretation of the implication of the biological metaphors in these two contexts in the next two sections respectively.

Kant links his philosophy of history to his theory of generation, by stating that man must be regarded as engaged in a transgenerational project of self-fulfillment, because of the fact that his natural capacities are indefinite and the fact that his life is definite, so that it is only as a species, and not as an individual, that he may be regarded as developing his inborn potential. This connection between the philosophy of history and the philosophy of biology has been regarded as an embarrassment by many recent commentators. Pauline Kleingeld (1999: 60), for instance, makes a point of the fact that we can no longer endorse Kant's pre-Darwinian teleological model. Allen Wood spelled out the reason for this insistence in the following words:

A more serious problem for Kant's philosophy of history is that we can no longer believe, for instance, in Kant's heuristically motivated natural teleology as the right way for investigating the structure and behavior of living organisms. Since Darwin, it has been recognized that the unconscious and unintended purposive arrangements in living things have a determinate empirical explanation based on natural selection. Moreover, this explanation reveals that Kant's heuristic assumption that the teleology in organisms is maximal is empirically and explainably false. When we learn how the organs of a living thing evolved, for instance, we sometimes come to understand why they are not optimally suited for the function they perform. And it might turn out, for similar reasons, that not everything we rightly conceptualize as one of the "species predispositions" of an organism would have to be fully developed in the normal course of the organism's development. The biological basis of Kant's philosophy of history therefore seems to have been undermined by scientific developments between his time and ours. (Wood 2006: 259-260)

Kleingeld and Wood of course do not want to say that Kant's theory of history is thereby completely discredited. Wood admits as much immediately after voicing his reticence towards Kantian biology:

Yet it is not so clear that the methodological considerations motivating Kant's philosophy of history are less applicable today than they were in the eighteenth century. Biology may have made advances that undermine the application to it of Kant's heuristically motivated natural teleology, but human history is still an area of inquiry to which no similar empirical theory has been applied with success. It may be that our best chance of making it intelligible is still the regulative-teleological one that Kant adopts. (Wood 2006: 260)

Scholars like Kleingeld and Wood are therefore forced to take an ambiguous stance towards Kant's philosophy of history, in which it is seen, on the hand, to rely on untenable biological assumptions, and on the other hand, to have the clear benefits of giving us a basis for theoretical inquiry into history and of harmonizing the noumenal and phenomenal aspects of man. The first problem it is believed to solve by stating that, although individual intentional actions by free individuals cannot be the subject of a science, the underlying non-intentional processes driving human history can provide a foothold for a scientific historiography. The second solution is considered to consist in the idea that, even though the predispositions of man, and therefore his rationality, is fixed, atemporal and universal, these predispositions can be regarded as requiring history to develop themselves.

Authors tend to diverge on which question they regard as essential. Allen Wood focuses on the problem of a theoretical approach to history and Alix Cohen (2007) has developed a richer account of how Kant can be read as giving a basis for not just historiography, but for the human sciences in general. For these authors, the problem is mostly how the freedom of individual intentional acts can be squared with the question of causes, laws or trends in history or in culture. As a result, this problem is more concerned with the status of the social sciences and the humanities.<sup>1</sup>

---

<sup>1</sup> I also suspect that Allen Wood may be interested in Kant's theory because it is the basis of the idea of a non-intentional course of history with an internal mechanism or driving force. This idea is important to all those interested in for instance the Marxian idea of historical materialism, in which history is approached as subject to general historical tendencies. Allen Wood himself made great efforts in trying to show how Marxian historical materialism can be compatible with the ideas of agency and free will. In his book on Karl Marx, he wrote the following on the topic: "Historical materialism seeks to explain political events or socially prevalent ideologies teleologically by showing how they contribute to basic social or historical tendencies. Explanations of this sort are in general not deterministic. They do not imply that the tendency in question had to be manifested in just that way, or that it required the contribution of that particular *explanandum* in order to exist at all. In fact, a 'basic historical tendency' which required (that is, depended on) certain particular events or the performance of certain actions by particular individuals would *eo ipso* be too flimsy and brittle to

Other authors are more concerned with the ethical dimension of Kant's thought. Here, the concern is that Kantian ethics is dependent on the universality of the moral law, which is believed to imply not only the universality of reason, but also that of the other faculties. If this is so, however, allowing for a truly historical and developmental aspect to the a priori would destroy this claim to universality. This is why many interpretations are so invested in proving that ultimately, there is little genetic to Kant's conception of genesis, little historical about his conception of history, little development to his conception of development. And this is precisely the conclusion that the talk of natural predispositions yields according to, for instance, Karl Ameriks:

It turns out that Kant's conception of his own view as epigenetic involves genesis or development only in a very restricted sense. As interpreters such as Zammito and Zoeller have argued in different ways, Kant's denial of specific versions of a preformation account of our ideas is still compatible with an emphasis on the strict a priori nature of some of them, and even with an allowance that we are dependent on something that is preformed. Even though Kant does not want to call any of our ideas innate, he does insist that our a priori ideas are very unlike empirical ones because they are "originally" rather than derivatively "acquired." Kant still takes the "ground" of the possibility of pure representations to be innate in us, although not the representations themselves, because he wants to avoid the odd notion of a kind of actual mental picture slumbering within the mind without any epistemological activity. (Ameriks 2009: 63)

Here, Ameriks is using the Sloan-Zammito interpretation of Kant's developmental theory as ultimately preformationist to deny that Kant's philosophy can allow for historicity at all, except through an equivocation of the word "historical". He is specifically doing so by referring to the fact that Sloan and Zammito have inferred, from Kant's purported allegiance to some form of preformationism, that Kant's epistemology does not allow for the dynamic and historical element that we now regard as a virtue of a theory of knowledge. There appears to be some disagreement over the degree of fixism to which Kant is committed by his alleged mitigated preformationism, but all authors do tend to close the gap between the Kantian and rationalist theories of knowledge.

---

deserve the name. A materialist explanation of the French Revolution, for example, might show that the events of 1789 and the actions of such men as Mirabeau and Sieyes served to bring about political changes required by the economic conditions of France and the state of the class struggle. But the explanation need not show that economic conditions required these changes to take place at just that time or in just that manner, and almost certainly could not show that they had to be effected by just those individuals. To say that the Revolution came about on account of a basic historical tendency is precisely *not* to say that it had to happen in just the particular way it did." (Wood 2004: 114)



In the following two sections, I will argue that Sloan and Zammito, and those authors building on their theses, are mistaken in minimizing the dynamism available to Kant. I will show that all the relevant passages can be read as in line with the epigenesis theory I ascribed to Kant in the previous two sections. I will also show that an epigenesis and dynamic reading makes better sense of Kant's criticisms of other philosophers than the preformationist and fixist reading because the latter makes it unclear how Kant differed from his preferred targets.

## 3.2 Epigenesis and Transcendental History

### 3.2.1 Rousseau's Pristine Man

One of the most famous biographical details of Kant's otherwise unremarkable personal life is his daily routine. Ever since Heinrich Heine's (1861: 186-187) evocative interpretation of the connection between this *petite histoire* and the meaning of Kantian philosophy, authors and teachers have relished in recounting the fact that Kant's daily routine, including his morning walk, was obsessively planned and consistently upheld. They equally like to point out the sole exception to Kant's observance of this daily routine. One day, so legend has it, Kant was so deeply immersed in his reading of Rousseau's *Emile ou de l'éducation* that he lost all awareness of his inflexible schedule. This legend contains two elements: the idea of a rule and a regularity governing Kant's whole being, and the idea of an exception to that rule that is, of course, a confirmation of the rule. But what does it show? It might be regarded as a solid reminder of the fact that Kant was so impressed and hence influenced by Rousseau that his love for the latter broke through his otherwise unswerving allegiance to his principles. But this is a problematic statement, for *interest* does not track *influence*, and *influence* does not mean *agreement*. Instead, many readers of Kant and Rousseau insist on the great chasm that divides the two authors. Kant may perhaps have been Rousseau's greatest critic, and nowhere is this so evident than in their respective views on Education.

Rousseau's *Emile* is a massive work on education, but not only on education. It collects some of the most poignant statements of Rousseau's philosophy and general commitments. One of its most disconcerting features is that this treatise on education is just as much a treatise *against* education. Rousseau believes to have observed that the men of his era are subject to two contrary kinds of impulses, those of nature and those of society, and that the resulting contradictions doom them to unhappiness:

From these contradictions is born the one we constantly experience between ourselves. Swept along contrary routes by nature and by men, forced to divide ourselves between these different impulses, we follow a composite impulse which leads us to neither one goal nor the other. Thus, in conflict and floating during the whole course of our life, we end it without having been good either for ourselves or for others. (Rousseau 1979: 41)

The goal of Rousseau's education is to raise man such as nature would have him develop. The hope is that a man raised primarily to be human will be more in harmony with himself, that he will become that noble creature which Rousseau calls natural man. The challenge of education is then the following: "To form this rare man, what do we have to do? Very much, doubtless. What must be done is to prevent anything from being done." (Rousseau 1979: 41). Rousseau's theory of education is a theory of counteracting education, of de-education. It is a theory of counteracting the adverse effects of socialization on our capacities and moral fibre, of restoring man to his original virtue. The opening line of the first book already announces this through the famous phrase: "Everything is good as it leaves the hands of the Author of things; everything degenerates in the hands of man." (Rousseau 1979: 37).

The meaning of this phrase only becomes clear in the second book of *Emile*, where Rousseau advances his theory of freedom and happiness, which is condensed into what he calls his fundamental maxim: "The truly free man wants only what he can do and does what he pleases" (Rousseau 1979: 84). This maxim expresses Rousseau's conviction that man is at his best when his faculties and his desires are in harmony. In the state of nature, so Rousseau believes, man does not desire anything that is beyond his faculties, beyond his power. Furthermore, he advances that "[o]ur happiness consists [...] in the disproportion between our desires and our faculties" and that "human wisdom or the road to happiness" consists in "diminishing the excess of the desires over the faculties and putting the power and the will in perfect equality" (Rousseau 1979: 80).

There are two aspects of Rousseau's theory as expounded in the *Emile* to which we need to draw attention at this point. The first aspect is that of faculty rigidity, namely the idea that man has a distinct and well-circumscribed set of faculties and desires that are naturally present in him and by which he needs to live. The second is the idea of decadence, which assumes that man can leave the state of nature and enter into a state in which the relation between faculties and desire becomes distorted to such an extent that the good life becomes all but impossible. In the remainder of this subsection, I will discuss these two features of Rousseau's theory, so that I can address Kant's responses to them in the following subsections.

In the *Emile*, Rousseau advances, or assumes, that man has a determinate place in nature. This place is determined by man's nature, i.e. the faculties of which he disposes as a solitary being in an uncultivated world. It is for this reason that Rousseau implores us to do the following:

O man, draw your existence up within yourself, and you will no longer be miserable. Remain in the place which nature assigns to you in the chain of being. Nothing will be able to make you leave it. Do not rebel against the hard law of necessity; and do not exhaust your strength by your will to resist that law—strength which heaven gave you not for extending or prolonging your existence but only for preserving it as heaven pleases and for as long as heaven pleases. Your freedom and your power extend only so far as your natural strength, and not beyond. All the rest is only slavery, illusion, and deception. (Rousseau 1979: 83)

The question raised by this advice is of course why Rousseau feels the need to make it. If man has a properly assigned place in nature, then how is he ever led to leave it? How *can* he even go beyond his natural state if this state determines him fully? Rousseau's answer to this is that man's *superbia* stems from the superfluousness of some of his faculties: "All the animals have exactly the faculties necessary to preserve themselves. Man alone has superfluous faculties" (Rousseau 1979: 81). The foremost superfluous faculty seems to be that of Imagination:

It is thus that nature, which does everything for the best, constitutes him [i.e. man] in the beginning. It gives him with immediacy only the desires necessary to his preservation and the faculties sufficient to satisfy them. It put all the others, as it were, in reserve in the depth of his soul, to be developed there when needed. Only in this original state are power and desire in equilibrium and man is not unhappy. As soon as his potential faculties are put in action, imagination, the most active of all, is awakened and outstrips them. It is imagination which extends for us the measure of the possible, whether for good or bad, and which consequently excites and nourishes the desires by the hope of satisfying them. But the object which at first appeared to be at hand flees more quickly than it can be pursued. It transforms and reveals itself in the distance ahead of us. No longer seeing the country we have already crossed, we count it for nothing; what remains to cross ceaselessly grows and extends. Thus one exhausts oneself without getting to the end, and the more one gains on enjoyment, the further happiness gets from us. (Rousseau 1979: 80-81)

This paragraph is, in my opinion, central to understanding the stance Kant will adopt against Rousseau's philosophy of education, philosophical anthropology and philosophy of history. It describes man as a creature that has a fixed set of faculties, some of which are latent at first and only develop later on in the history of the individual and/or the species. One of these faculties is the ability to conceive possibilities that lie beyond one's reach, and to conceive of possible future states. As Rousseau laments: "Foresight! Foresight, which takes us ceaselessly beyond ourselves and often places us where we shall never arrive. This is the true source of all our miseries." (Rousseau 1979: 82) The most important illusion to which Imagination leads is the idea of extending one's power, of transcending an initial state which is considered to be one of weakness. Against this

idea, Rousseau insists that the “word *weak* indicates a relation, a relation obtaining within the being to which one applies it. [...] He whose needs surpass his strength, be he an elephant or a lion, be he a conqueror or a hero, be he a god, is a weak being.” For Rousseau, then, power expresses a relation between the faculties and the desires within a certain being. It is for this reason that he believes us to become weaker as we strive to go beyond our natural state:

Man is very strong when he is contented with being what he is; he is very weak when he wants to raise himself above humanity. Therefore, do not fancy that in extending your faculties you extend your strength. On the contrary, you diminish your strength if your pride is extended farther than it. (Rousseau 1979: 81)

What Rousseau warns against is the idle project of extending the faculties. This must necessarily fail, because the power of our faculties is limited, whereas the scope of the imagination is infinite.<sup>2</sup> I think it likely that Rousseau believes our powers to be rigidly fixed, but it is also possible that he simply stresses our powers cannot possibly increment proportionately to our desires. On this latter reading, Rousseau would not say that striving to extend our faculties is harmful because our faculties cannot be extended, but rather that this striving is ill-advised because it would always lead us to be dissatisfied with any given expansion of our faculties. The two readings can converge if we add the further thesis, which Rousseau seems to hold as well, that the reason why any expansion of our faculties will prove dissatisfactory is because the increment of our faculties is limited in a way in which the increment of our desires is not. The latter thesis is committed to clear preestablished limits on the faculties.

It appears, then, that Rousseau is committed to the following two doctrines, namely the fixity of our faculties, and the idea that there is in man a capacity for decadence. This decadence is the capacity to exit the state of nature and enter the state of culture,

---

<sup>2</sup> The idea of the mismatch, the disproportion, between our faculties and our imagination, is an echo of the common theme of the disproportion between the will and the understanding. In Descartes, and many authors after him, this disproportion has been used to explain the possibility of error and sin in man. The problem with the capacity to err is that it is incomprehensible why a just God would choose to give us a capacity specifically to err, and thereby possibly, by sinning, incur his wrath. The answer to this conundrum is that the capacity to err is not a real, positive capacity, but the result of the disproportion between will and intellect, the former being unlimited in scope whilst the latter is finite and definite. Thus, although we are always capable of knowing the true and doing the good, we at times find ourselves seduced by the possibility of the alternatives to the true and the good. Rousseau’s thesis that man can abstain from erring by residing within the bounds of his faculties, is a secularized version of this theory. The form in which Rousseau uses it is related to the Lockean idea of empiricist humility, which insists that we keep within the scope of the faculties, since these are unproblematic as long as we use them in the way God or nature intended us to use them. We will see in the rest of this dissertation that Kant’s so-called “humility” is not clearly related to this scheme.

enter into society. This passage from nature to culture is, for Rousseau, always one of decadence, one of loss rather than gain:

Society has made man weaker not only in taking from him the right he had over his own strength but, above all, in making his strength insufficient for him. That is why his desires are multiplied along with his weakness, and that is what constitutes the weakness of childhood as compared to manhood. [...] We were made to be men; laws and society have plunged us once more into childhood. (Rousseau 1979: 85)

For Rousseau, culture is vanity, vanity in the sense of mere pride, hubris, a false feeling of power, and vanity in that it is built on emptiness, on vain, impossible desires and wishes. If man wants to be happy, he should remain in his original state, and constantly fight his ingrained urge to go beyond it.

### 3.2.2 Kant's Pedagogy

It is hardly controversial or original to stress that Kant's philosophy of pedagogy, his philosophy of history and his philosophical anthropology are to be related to and contrasted with Rousseau's. However, this mundane historiographical insight does not in itself yield a precise picture of what exactly Kant took issue with in Rousseau, and what the broader implications of the differences between their respective philosophies turn out to be. In the remainder of this section, I will advance the thesis that Kant was primarily interested in dispelling what I will call the Myth of the Origin, i.e. the myth that man can be considered as having an original state that later becomes altered or developed through history. I will argue, more specifically, that Kant's strategy in arguing against the Myth of the Origin consists in staging the conflict between the two core doctrines of Rousseau that I have identified in the previous section. In this subsection, I will show how this is done in the writings on pedagogy and anthropology, whereas in the next sections I will focus on the writings on history. In the end, however, I believe that this distinction is very artificial indeed, since these writings all advance the same point in different ways.

Recent interpreters have drawn more attention to the fact that Kant employs the talk of germs and predispositions, of *Keime* and *Anlage*, in his philosophical anthropology as much as in his philosophy of biology. Indeed, as we have seen, these terms were largely elaborated in Kant's writings on the very idea of human races. Now, we find Kant putting them to use in his theory of education:

Many germs lie within humanity, and now it is our business to develop the natural predispositions proportionally and to unfold humanity from its germs and to make it happen that the human being reaches his vocation. Animals fulfill their

vocation automatically and unknowingly. The human being must first seek to reach his, but this cannot happen if he does not even have a concept of his vocation. It is also completely impossible for the individual to reach the vocation. Let us assume a fully formed first human couple, and let us see how they educate their pupils. The first parents already give the children an example which the latter imitate, and that way some natural predispositions are developed. But not all predispositions can be developed in this manner, for the children only see these examples in occasional circumstances. Formerly, human beings did not even have a conception of the perfection which human nature can reach. We ourselves are not even yet clear about this concept. But this much is certain, that individual human beings, no matter what degree of formation they are able to bring to their pupils, cannot make it happen that they reach their vocation. Not individual human beings, but rather the human species, shall get there. (AA IX: 445-446)

At first glance, this paragraph reveals Kant to be committed to a doctrine also held by Rousseau, namely that of faculty fixism. In the previous subsection, I argued that Rousseau believed man to have a fixed set of capacities, and that these capacities determine his place in nature. These faculties cannot be truly altered or expanded, although some may be latent at first. Kant's talk of germs and predispositions seems to invoke the same picture here, since it seems to commit him to the idea that man has a prefixed nature, which may be merely latently present. The major difference between Kant and Rousseau is then that Kant believes it to be desirable that we develop our capacities, whereas Rousseau believes this would be detrimental to the individual and to the species.

Here, however, as often, the first glance is deceiving. In fact, the one divergence between the two authors that I have indicated ends up affecting their entire theories, such that what seem agreements between Kant and Rousseau are actually their fundamental differences. In this passage, Kant speaks not of the man's place in nature, as Rousseau would have, but of man's *vocation*. Both thinkers offer us the maxim that man must, above all, be human. For Rousseau this means that man must remain in his allotted place, and not seek to distort his original nature. For Kant, however, this means that man is never at place in nature, but is always that which is seeking his place in nature. For Rousseau, man is a location, whereas for Kant he is a vocation.

This is a first reason why Kant adopts the language of germs and predispositions. For Rousseau, man's essential faculties are present in his original state. For Kant, they are never truly actualized, and always to be actualized through the process of education. Education and culture lead us to develop these germs and predispositions that lie hidden in us. But we should be careful not to conclude from this that the germs and predispositions are nevertheless on a par with man's capacities as they are conceived in Rousseau's theory. The crucial difference lies in the fact that, according to Kant, the process of education is not just one of seeking to occupy one's intended place in nature

– it is equally the process of seeking out what one’s place in nature might be. This is why Kant stresses that man does not know his vocation, which means that he does not know his germs and predispositions either, since these only reveal man’s place in nature as they reveal themselves:

The human being can only become human through education. Man is nothing except what education makes out of him. It must be noted that the human being is educated only by human beings, human beings who likewise have been educated. That is also why the lack of discipline and instruction in some people makes them in turn bad educators of their pupils. If some day a being of higher kind were to look after our education, then one would see what the human being could become. But since education partly teaches the human being something and partly merely develops something within him, one can never know how far his natural dispositions reach. (AA IX: 444)

Here, Kant identifies part of the problem as our incapacity to ever decide exactly which the nature of man is, or which the vocation of man is, and our irremediable ignorance of which our natural predispositions are.<sup>3</sup> This means that his talk of predispositions is not meant to indicate the innate, the ingrained capacities of man, but rather to indicate that we must understand man as having an essence, although this essence is not fixed. The thesis that “man is nothing except what education makes out of him” is tantamount to the idea that there can be no meaningful talk of man as a specific species with specific limits. The limits are assumed to be there, but we can never make determinate claims about them. Kant’s humanism is thus peculiar for not essentializing the nature of humanity.

But this raises questions of its own, most importantly as to why Kant chooses to employ this talk of germs and predispositions. Why would he feel the need to indicate that man has ingrained capacities if he does not believe that these capacities are fixed? What are we to understand by non-fixed innateness? I believe this problem can be solved by closely considering the problem of what it means to see man as that paradoxically natural being that somehow only exists as non-natural. Why the status of

---

<sup>3</sup> One might object that, in this passage, Kant states that a perfect creature *could* reveal what could be made of man, and that this means that is only accidentally the case that our essence is unknown (namely because we have only fallible teachers). But this is dependent on the assumption that he takes this perfect creature in a positive sense, and not a negative one. Kant, however, almost always invokes the idea of perfect being as a negative concept. On such a plausible negative reading, the caveat might actually mean the following. Man’s nature is potentially infinitely malleable. Since human beings are in principle finite, they could never realize this infinite potential. If there were a creature that is capable of actualizing the infinite, i.e. God, he would be able to realize and oversee the whole of man’s capacities. But this is not the situation in which we find ourselves.

man as a natural being is paradoxical will become clear when we turn our attention to the tension between Rousseau's theory of original man and his theory of decadence.

Rousseau assumes that man is an animal like other animals, even though he is somehow distinguished by freedom. It would be wise for man to keep this in mind at all times, for it would indicate how his faculties are somehow appropriate for a specific kind of life, namely the solitary life of the savage. His faculties are not intended to function in the artificial context of society and culture, which will therefore end up corrupting him.<sup>4</sup> This theory requires Rousseau to explain, however, why it is that man tends to exit this original state, and how man can ever exit his circumscribed space. Rousseau does so by indicating that there is a natural capacity in man, namely imagination, which allows him to conceive of a state other than his current one. But this raises the further question as to whether this capacity is natural or non-natural, and consequently whether the artificial state is a natural one or a non-natural one. If the capacity for imagination is a natural one, then any employment of it will be natural, and will not take man out of nature. Society will be but one way in which man is in nature, i.e. but one way in which man can occupy his place in nature. If this is true, it is not clear how Rousseau's admonition to keep to our nature would be able to withhold us from societal life. In order for us to be capable of exiting nature through the employment of imagination, imagination must already be non-natural to man, must already be decadent. But then it must either be due to an earlier decadence, an earlier fall from a natural state, which would then also either be due to a natural or a non-natural faculty, or the natural state of man would already be decadent. I believe Kant saw this problem, and decided therefore that man, as a natural being, is that being that is necessarily non-natural. His fundamental faculties are those which automatically remove him from the natural order; hence the only properties essential to man are those which preclude him from having essential properties.

What are these properties? As the second proposition of the *Idea for a Universal History With a Cosmopolitan Aim* reveals, it is reason:

*In the human being (as the only rational creature on earth), those predispositions whose goal is the use of his reason were to develop completely in the species, but not in the individual. Reason in a creature is a faculty of extending the rules and aims of the use of all its powers far beyond natural instinct, and knows no boundaries to its projects. But reason itself does not operate instinctively, but rather needs*

---

<sup>4</sup> It is interesting to note that these ideas are in a way held by contemporary (ultra)darwinist naturalists as well, since they insist that our cognition evolved in a situation, or series of situations, in which the selective pressures were entirely different from the ones in which we tend to live since the end of our real evolution, i.e. cultural states. Perhaps Kant's criticism of Rousseau will serve equally well to dispel the many non sequiturs that invoke the name of Darwin as an argument from authority.



attempts, practice, and instruction in order gradually to progress from one stage of insight to another. Hence every human being would have to live exceedingly long in order to learn how he is to make a complete use of all his natural predispositions; or if nature has only set the term of his life as short (as has actually happened), then nature perhaps needs an immense series of generations, each of which transmits its enlightenment to the next, in order finally to propel its germs in our species to that stage of development which is completely suited to its aim. And this point in time must be, at least in the idea of the human being the goal of his endeavors, because otherwise the natural predispositions would have to be regarded for the most part as in vain and purposeless; which would remove all practical principles and thereby bring nature, whose wisdom in the judgment of all remaining arrangements must otherwise serve as a principle, under the suspicion that in the case of the human being alone it is a childish play. (AA VIII: 18-19)

This passage presents a great number of interpretative difficulties. In the following subsection, I will attempt to provide a better interpretative access through a sustained reading of another major, though regularly overlooked, text by Kant on the philosophy of history from the 1780s. There, it will become clear that we need to understand reason as the capacity for history, rather than as an ahistorical capacity.

### 3.2.3 The Myth of the Origin

In the second section, I already briefly advanced the idea that Kant *staged* the contradiction between Rousseau's two core doctrines, namely that of natural man and that of the decadence of man. It is this staging that has given the impression to some (e.g. Philonenko 1972: 43-50) that Kant is often misappropriating or misinterpreting Rousseau's theory in order to let the latter agree with him. In this section, I will show that Kant was instead concerned with internally subverting some of the core assumptions behind the very theories he seems to appropriate, namely the preformationist theory of development and the idea of natural, non-decadent man.

In order to understand the possibility of an alternative interpretation of the notion of a natural predisposition in Kant's philosophy of history, I believe we should turn to his text *Conjectural Beginning of Human History*, from 1786. In a way, this text serves to complete the 1784 essay *Idea for a Universal History with a Cosmopolitan Aim*, most importantly the second proposition, in which Kant speaks of natural predispositions. The title already suggests that Kant regards his account in this paper as merely conjectural. Yet, he does insist that his conjecture here is not on a par with the conjectures that one may make in historiography in order to fill in the gaps of the historical record. The reason is that his question here is not that concerning the history of man, i.e. of the development of freedom and reason, but that concerning the "first

development of freedom from its original predisposition in the nature of the human being” (AA VIII: 109).

The most peculiar aspect of Kant’s conjectural history is probably the fact that it chooses to roughly follow the strand of the Biblical Myth of the Fall in the book of Genesis. Some authors, particularly Arthur Lovejoy (1959: 193-195), have taken issue with this choice, believing that it reveals Kant’s enduring commitment to the Biblical narrative of history. In Lovejoy’s case, this serves as yet another argument to deny the historical and the transformative in Kant’s account, since the book of Genesis precludes a Darwinian version of the origin of species mainly through selective pressures on biopopulations in which minor inheritable differences between individuals occur. Lovejoy is right to claim that the Kantian picture is not very conducive to an evolutionary picture of the place of man in nature, but I believe it is wrong to therefore dismiss it as merely a testimony to Kant’s lingering commitment to the Christian world-view.

What Lovejoy and others overlook, in my opinion, is the fact that the Myth of the Fall has always been interesting to philosophers for the philosophical idea that lies behind it, not (just) for its literal properties. One of the best explicitations of this philosophical content has been offered by Hegel in his *Encyclopaedia of Philosophical Sciences*. In one of the *Zusätze* of the *Encyclopaedia Logic*, he states the following:

It appears appropriate to consider the myth of the Fall at the very beginning of the Logic, because the Logic is concerned with cognition, and the myth too deals with cognition, with its origin and significance. Philosophy should not shy away from religion, and adopt the attitude that it must be content if religion simply tolerates it. And, on the other hand, we must equally reject the view that myths and religious accounts of this kind are something obsolete, for they have been venerated for millenia by the peoples of the world. (Hegel 1991: 61)

Hegel thus makes the point that the Myth of the Fall is not just an article of faith to be held by the religious, nor is it just a historical narrative of an event in our eschatological history. He insists instead that it is really about something more general, i.e. cognition:

When we are comparing the various forms of cognition with one another, it can easily appear that the first one, that of immediate knowledge, is the most adequate, the finest, and the highest. It includes everything that is called innocence in the moral sphere, as well as religious feeling, simple trust, love, fidelity, and natural faith. Both of the other forms, first reflective cognition and then philosophical cognition too, step out of that immediate natural unity. Insofar as they have this in common with one another, the mode of cognition that consists in wanting to grasp the truth through thinking can easily appear as the human conceit that wants to recognise the true by its own strength. This standpoint of universal separation can certainly be looked on as the origin of all wickedness and evil, as the original transgression; and on this view it seems that

thinking and cognition must be given up in order to return [to unity] and become reconciled again. As for the abandonment of natural unity here, this marvellous inward schism of the spiritual has been something of which all peoples from time immemorial have been conscious. An inner schism like this does not occur in nature, and natural things do not do evil. We have an old account of the origin and consequences of this schism in the Mosaic myth of the Fall. The content of this myth forms the foundation of an essential doctrine of faith, the doctrine of the natural sinfulness of man, and his need of help to overcome it. (Hegel 1991: 61)

Hegel argues that the Myth of the Fall resonates with our tendency to believe that an initial state of immediacy is somehow possible and desirable, and that mediation is a form of decadence and the first ground of evil, the “original transgression”. There is also the belief that the root of this evil is “conceit” in the form of an attempt to use our reason. Again, there is the idea that “[a]n inner schism like this does not occur in nature, and natural things do not do evil”, namely that it is somehow testimony to the unnaturalness of man that he is capable of evil and sin. And lastly, there is the following peculiar aspect to the theory:

In our Mosaic myth, moreover, we find that the occasion for stepping out of the unity [of innocence] was provided for humanity by external instigation (by the serpent). But in fact, the entry into the antithesis, the awakening of consciousness, lies within human beings themselves, and this is the story that repeats itself in every human being. (Hegel 1991: 62)

What Hegel insists on is that, in the Biblical narrative, the serpent is not really the cause of the Fall, but merely its occasion. The Fall is essential to man, since his nature is determined by a certain non-naturalness, that of reason, which will inevitably lead him outside the original state of happiness.

Hegel’s philosophical analysis of the Myth of the Fall should remind any reader of Rousseau’s account of original man. Rousseau too postulates an original state in which man is happy and at one with nature. He too believes no creature can naturally go beyond its proper station where it is happy. He too believes that man has a peculiar faculty, in his case imagination, which causes him to develop vain wishes, and believes that these vain wishes lead him out of his original station and into a state of error and unhappiness. It is Rousseau’s narrative that is complicit with the Biblical tale, and revealing of its philosophical rationale.

I believe Kant was perfectly aware of this, and that he chose the Biblical narrative as a template not to get the approval of the religious, nor because of his own sincere religious beliefs, but because it can serve to unveil and deconstruct Rousseau’s ideas of origin and decadence. This is why Kant deems it important to explicate the setting of the book Genesis as constraints necessary for a kind of thought experiment (AA VIII: 110). First, the assumption that we are dealing with the first pair of humans is necessary

in order to account for the fact that they are not already socialized through their life with siblings and parents. By making this abstraction, Kant is already revealing that he deems it unlikely that there was ever a state at which man lived solitary and not already in social context, albeit a crude one. Second, the assumption that we are dealing with adults who have already received certain capacities is meant to waver further conjectures on how they came to acquire these. They are of course also meant to meet Rousseau's picture that there are certain capacities that already develop in the state of nature and are not themselves detrimental to man. The very setting is thus not meant to approach the Biblical situation, but to show that the Biblical situation shares essential properties with Rousseau's state of nature.

Then Kant specifies that life in his counterfactual garden is natural life because in it "[i]n instinct, that *voice of God* which all animals obey, must alone have guided the novice" (AA VIII: 111). It is easy to read Kant as defining instinct as a divine command. I believe it is more likely that he is showing that the divine command plays the role in the Biblical narrative that instinct plays in the conjectures of more naturalistically inclined theorists, namely that of determining the original constraints of the human being.

Having stipulated his original state in such a way that the parallels between Rousseau's state of nature and the Mosaic Garden of Eden are maximally visible, Kant goes on to discuss the fundamental event which we need to explain:

As long as the inexperienced human being obeyed this call of nature, he did well for himself. Yet *reason* soon began to stir and sought through comparison of that which gratified with that which was represented to him by another sense than the one to which instinct was bound, such as the sense of sight, as similar to what previously was gratifying, to extend his knowledge of the means of his nourishment beyond the limits of instinct (*Genesis* 3:6). This attempt might have happened to turn out well enough, although instinct did not recommend it, if only it did not contradict it. Yet it is a property of reason that with the assistance of the power of imagination it can concoct desires not only *without* a natural drive directed to them but even *contrary* to it, which desires in the beginning receive the name of concupiscence, but through them are hatched bit by bit, under the term voluptuousness, a whole swarm of dispensable inclinations, which are even contrary to nature. The occasion for deserting the natural drive might have been only something trivial; yet the success of the first attempt, namely of becoming conscious of one's reason as a faculty that can extend itself beyond the limits within which all animals are held, was very important and decisive for his way of living. (AA VIII: 111-112)

We should pause to ponder the various implications of this passage. First of all, Kant already assumes that man has the possibility of reason. This is not an outrageous assumption for those targeting Rousseau, since the latter was indeed committed to the fact that man does dispose of a faculty that allows him to conceive of possible

alternatives. Furthermore, Kant assigns a function in this process of the Fall to imagination, the faculty singled out by Rousseau as the malicious counselor whispering evil thoughts into our ears. Of course the difference here is that Kant wants to give standing to the faculty for leaving the state of nature by calling it reason rather than imagination, whereas Rousseau gives it the name of the faculty that has met with opprobrium and even hostility in many early modern philosophies.

Secondly, Kant points out that reason and imagination are responsible for devising unnatural desires. Here, unnatural means both counter to instinct and counter to divine command. These desires are explained through the same mechanism as that used to explain the vain desires in Rousseau's *Emile*, namely the consideration of unwise alternative ways of life, of alternative possibilities that we were never meant to entertain.

Thirdly, and most importantly, Kant is insistent that the Fall itself would not have been a particularly great event. It does not require a particularly great upheaval or betrayal to wrench man from his natural, original state. On the contrary, the smallest of successes would have set us on the road of culture, would have dislodged our natural self-sufficiency. This is important for two reasons. First of all, it speaks against the Christian doctrine that the Fall was due to a great transgression on our part, having broken the one clear commandment we were given. Second, it challenges the idea that we need an external trigger to commit this first and ultimate crime; much like Hegel insisted that the serpent's seduction was not specifically required as a cause since we bore within ourselves the fruits of sin, so Kant insists that any small event can provide the occasion for the great transgression, since what matters is that the root lies within us, not that it is somehow occasioned from without:

Now the harm might have been as insignificant as you like, yet about this it opened the human being's eyes (*Genesis 3:7*). He discovered in himself a faculty of choosing for himself a way of living and not being bound to a single one, as other animals are. Yet upon the momentary delight that this marked superiority might have awakened in him, anxiety and fright must have followed right away, concerning how he, who still did not know the hidden properties and remote effects of any thing, should deal with this newly discovered faculty. He stood, as it were, on the brink of an abyss; for instead of the single objects of his desire to which instinct had up to now directed him, there opened up an infinity of them, and he did not know how to relate to the choice between the; and from this estate of freedom, once he had tasted it, it was nevertheless wholly impossible for him to turn back again to that of servitude (under the dominion of instinct). (AA VIII: 112)

The reason why the discovery of reason is so transformative an experience is that it is also the discovery of freedom. With this discovery, Kant insists, we would have once and for all left the domain of nature, left the possibility of guidance by instincts. This new,

qualitative switch may be accompanied by a kind of vertigo, as we leave the realm of the definite and enter that of the indefinite. But this feeling of disconcertedness may very well equally be a feeling of the sublime, of our capacity grasp the infinity of freedom.<sup>5</sup> What is certain is that there is no going back: the realm of freedom has become second nature to man.

But is it only second nature? This is the problem Kant's analysis pushes on us, for it reveals that the Fall was necessary, and was grounded in one of the fundamental capacities of man, namely that of *reason*. It is our nature as *rational animals* that have destined us to never be a definite kind of animal; for it is our reason that prompts us to regard ourselves only under the aspect of possibility. It is essential to man *that* he commit the original sin, *that* he leave the original state, since his essential trait is precisely his inability to remain there. This is Kant's crucial shift with respect to Rousseau, namely that the determination of the natural predispositions is meant to show how we are always destined to forever make shape our own faculties. That this whole idea is meant to solve the problems facing Rousseau is explicit in the *Conjectural Beginning* itself:

In this manner one can also bring into agreement with themselves and with reason the assertions of the famous *J.-J. Rousseau*, which are often misinterpreted and to all appearance in conflict with one another. In his writing *on the influence of the sciences* and *on the inequality of human beings*, he shows quite correctly the unavoidable conflict of culture with the nature of the human species as a *physical* species in which each individual was entirely to reach his vocation; but in his *Emile*, his *Social Contract* and other writings, he seeks again to solve the harder problem of how culture must proceed in order properly to develop the predispositions of humanity as a *moral* species to their vocation, so that the latter no longer conflict with humanity as a natural species. From this conflict (since culture, according to true principles of *education* of human being and citizen, has perhaps not yet rightly begun, much less having been completed) arise all true ills that oppress human life, and vices that dishonor it; nevertheless, the incitements to the latter, which one blames for them, are in themselves good and purposive as natural predispositions, but these predispositions, since they were aimed at the merely natural condition, suffer injury from progressing culture and injure culture in turn, until perfect art again becomes nature, which is the ultimate goal of the moral vocation of the human species. (AA VIII: 117-188)

Rousseau's problem is, then, for Kant, that he postulates a dualism of nature and culture, of the realm of instinct and the realm of reason, such that the only correct response would be to reject culture or mitigate its effects and live within our natural

---

<sup>5</sup> On this, see further chapter 7.

bounds. For Kant, the idea is rather to have our cultural, moral, rational self become our nature.<sup>6</sup> The original dispositions of man have one and only one goal: to propel man from his original, natural state into an uncertain future that is of his own making.

All this is yielded by only one decision on Kant's part, however, namely to regard *all* tendencies of man as equally useful. Remember that the core of Rousseau's theory is that man is naturally born with superfluous capacities. It is these superfluous, unnatural elements of our nature that are the cause of our alienation from nature. Kant decides that it does not make sense to define a point of origin together with a principle of departure from this origin without acknowledging that this principle is included in the origin. It makes no sense to think of reason or imagination as an originary supplement. All that it reveals is that, in reflecting on origins, we want to specify their nature without acknowledging the fact that, because of the fact that the principle of decadence is already included in it, the origin has always already passed. It is always a post facto mythical reconstruction we use in order to understand our own heading, which is never clear to us in its uncertainty.

This is in my opinion the real meaning of Kant's talk of natural predispositions: they are the necessary assumptions we must make in order to understand our own history as teleological, as being guided by some principle, and in order to understand how we could have come to have the unnatural nature that we have. The natural predispositions are the assumed original correlates of our current capacities, the reasons for our current situation. The progress of man can be regarded in two ways: as an idea of reason drawing us to the end of our history, or as a *vis a tergo*, a gust of wind blowing us out of paradise as we, with our backs to an unknown future, can see nothing but the ruins of the past amassing at our fleeting feet. For Kant, these two perspectives are equally necessary, and Rousseau's causal *Unheilsgeschichte* is the chiral of the Enlightenment idea of progress. In the Enlightenment project, we can understand ourselves as underway to something new, something greater, only if we already acknowledge that the possibility for this resides in us. And these possibilities reside in our natural predispositions, that will forever remain unknown to us. It is in this way that Kant's epigenesis opens the path to the idea of transcendental history, which we will further explore in section 3.4. First, however, we need to investigate the implications of Kant's embryological language in epistemology.

---

<sup>6</sup> The phrase "until perfect art again becomes nature" is due to Abbé Jean Terrasson. In fact, the dictum is to be found in the same chapter of the same work where Terrasson offers the idea that many books would have been much shorter, had they not been so short. It is thus more than likely that Kant picked up both ideas from the 1756 German translation of the 1754 book *La Philosophie applicable à tous les objets de l'esprit et de la raison*.

## 3.3 Epigenesis and the Transcendental History of Reason

### 3.3.1 Kant's Epistemological Preformationism?

In the *Critique of Pure Reason*, Kant refers to theories of generation on three separate occasions, each time to characterize a position in epistemology. Two of these appear in both editions, whereas the third makes its appearance only in the second edition of 1787. The puzzle offered by these various usages of embryological analogies is not only the difficulty in determining what exactly they are intended to convey, but also that the analogies used in the first edition seem to invoke preformationism, whereas those of the second edition refer approvingly to epigenesis. This has led Sloan and Zammito to claim that Kant's thinking changed somewhere around the mid-1780s.

The first important passage is from the opening of the Transcendental Analytic:

I understand by an analytic of concepts not their analysis, or the usual procedure of philosophical investigations, that of analyzing the content of concepts that present themselves and bringing them to distinctness, but rather the much less frequently attempted analysis of the faculty of understanding itself, in order to research the possibility of a priori concepts by seeking them only in the understanding as their birthplace and analyzing its pure use in general; for this is the proper business of a transcendental philosophy; the rest is the logical treatment of concepts in philosophy in general. We will therefore pursue the pure concepts into their first seeds [*Keime*] and predispositions [*Anlage*] in the human understanding, where they lie ready, until with the opportunity of experience they are finally developed and exhibited in their clarity by the very same understanding, liberated from the empirical conditions attaching to them. (CPR A65-66 / B90-91)

Zammito offers the following interpretation of this passage, drawing lavishly on Sloan's analyses:

the 1781 language is unequivocally a *preformationist* analogy. The concepts lie 'predisposed' in the understanding; they are not produced, they are occasioned. As Sloan argues, in terms of the philosophical debate about Kant's relation to innate ideas, this is clearly as 'nativist' a Kant as one can find. Moreover, Kant meant to suggest an element in the analogy which would be central to his thinking throughout, namely that just as *Keime* and *Anlagen* were inaccessible to ultimate derivation, so too the concepts of the understanding were simply givens behind which we could not seek. (Zammito 2003: 84)



I believe this interpretation is mistaken, because 1) Kant's language is not as unequivocally preformationist as Zammito assumes, and 2) it reads Kant as suggesting a nativism that he himself explicitly repudiated.

Firstly, Sloan and Zammito base their assumption that these passages are unequivocally preformationist on the assumption that Kant's usage of the terms germs and predispositions is unequivocally preformationist. In the previous chapter, I have offered reasons to believe that the latter assumption is false: Kant's usage of these terms does not square with classical preformationism, and seems intended rather to clarify his own position on the generative force. Contrary to what the preformationist reading suggests, Kant is saying here that the pure concepts of the understanding are pregiven only in so far as they constitute a constraint on the production of concepts. Zammito's suggestion that "the concepts of the understanding were simply givens behind which we could not seek" is misleading, then, because although the concepts might be givens, the specific form they take due to the occasioning experience is not. When taken seriously, the analogy with the so-called Kantian preformationism of the 1770s suggests that the pure concepts are not completely fixed, and can develop differently under different circumstances.

What Kant does stress is that we must take the specific forms that the concepts might take to be due not to influence from the outside, but as deriving from internal reasons. This too resonates with the 1770s-picture: the constraints on the generative force must be regarded as internal to this force, not as externally induced. In these passages, Kant therefore seems mainly concerned with showing that the concepts of the understanding are adaptable, but that this adaptability is not due to the direct influence of experience. This message is certainly different from the fixist and nativist reading that is often imposed on it.

There is another reason why we should be dissatisfied with the reading offered by Sloan and Zammito – a reason noted by Sloan himself in a footnote:

The absence of the *Keim-Anlage* language in the letter to Marcus Herz of February 21, 1772 in his important early discussion of the concept of the categories suggests that Kant developed his views more deeply on the generation question after this date, and possibly as a solution to the problem posed in the letter to Herz on how the categories could be brought into conformity with things. In this letter he rejects the thesis of Christian Crusius concerning "gewisse eingepflanzte Regeln zu urtheilen und Begriffe, die Gott schon so wie sie seyn müssen, um mit den Dingen zu Harmoniren, in die Menschliche Seelen pflanzte," but this comment follows his claim that the categories must still be grounded "in der Natur der Seele" (AK 10: 125–6). As I would interpret this letter in relation to the argument of this paper, the *Keim-Anlage* theory, in the unusual way Kant formulates this in 1775–77, allowed him to solve this problem by claiming that the *Keime* are pre-existent and determinate structures within the soul, but they do

not stand in a pre-established harmony with objects of experience, and they are only brought into play when activated by experience and the action of *Naturanlagen*. This avoids the problems of the pre-established harmony that he interprets as implied in Crusius's position. (Sloan 2002: 238)

In this footnote, Sloan brings up, and subsequently dismisses, a major objection to his reading, namely that Kant had already rejected nativism concerning the pure concepts of the understanding in the letter to Marcus Herz, and did not seem to believe that this discounted his preformationist analogy in the first Critique. His solution for this objection is unsatisfying, in my opinion, because the revised nativism that he prescribes to Kant is not revised at all. In fact, it seems to be precisely the position Leibniz endorses in the *New Essays on Human Understanding*. In that work, Leibniz had his mouthpiece Theophilus advance the following rebuttal of Locke's attack on the notion of innate ideas:

Since an item of acquired knowledge can be hidden there by the memory, as you admit that it can, why could not nature also hide there an item of unacquired knowledge? Must a self-knowing substance have, straight away, actual knowledge of everything which belongs to its nature? Cannot – and should not – substance like our soul have various properties and states which could not at all be thought about straight away or all at once? [...]

The mind is capable not merely of knowing them [i.e. the necessary truths], but also of finding them within itself. If all it had was the mere capacity to receive those items of knowledge – a passive power to do so, as indeterminate as the power of wax to receive shapes or of a blank page to receive words – it would not be the source of necessary truths, as I have just shown that it is. For it cannot be denied that the senses are inadequate to show their necessity, and that therefore the mind has a disposition (as much active as passive) to draw them from its own depths; though the senses are necessary to give the mind the opportunity and the attention for this, and to direct it towards certain necessary truths rather than others. [...]

So it [i.e. understanding] is not a bare faculty, consisting in a mere possibility of understanding those truths [i.e. eternal and necessary truths]: it is rather a disposition, an aptitude, a preformation, which determines our soul and brings it about that they are derivable from it. (G V: 75-77; Leibniz 1996: 78-80)

A careful reading of this passage shows that it offers precisely the view that Sloan ascribes to Kant in the first Critique. It does not rely on pre-established harmony at this point, nor is Leibniz committed here to the idea that *all* ideas are preformed in the mind in this manner. It does claim that the concepts and principles of the understanding that allow us to think necessity are and must be innate, and that only their development is occasioned by experience. Leibniz explicitly makes the analogy between this view and

preformation<sup>7</sup>. Should this lead us to believe that Kant has adopted this position in the first Critique?

Another passage from the first edition of the *Critique of Pure Reason* has prompted some to ascribe an anti-historical Leibnizianism to Kant, namely the following comment from the “architectonic of reason”:

Nobody attempts to establish a science without grounding it on an idea. But in its elaboration the schema, indeed even the definition of the science which is given right at the outset, seldom corresponds to the idea; for this lies in reason like a seed, all of whose parts still lie very involuted and are hardly recognizable even under microscopic observation. For this reason sciences, since they have all been thought out from the viewpoint of a certain general interest, must not be explained and determined in accordance with the description given by their founder, but rather in accordance with the idea, grounded in reason itself, of the natural unity of the parts that have been brought together. For the founder and even his most recent successors often fumble around with an idea that they have not even made distinct to themselves and that therefore cannot determine the special content, the articulation (systematic unity) and boundaries of the science. (CPR: A 834 / B 862)

At first sight, this passage is a good candidate for Zammito’s characterization of “unequivocally preformationist”. This is also why Yirmiyahu Yovel offers the following gloss on the element of history in the “architectonic of reason”:

The history of philosophy is thus a latent totality, governed by the tension between the inherent architectonic of reason and its particular stage of manifestation. All philosophical systems expound the self-development of reason, from which they get their organizing schema, and they are thus all united in a single organic whole. (Yovel 1980: 228)

Ironically, however, Yovel offers this as a gloss on the alinea following the one quoted above: “The systems seem to have been formed, like maggots, by a *generatio aequivoca* from the mere confluence of aggregated concepts, garbled at first but complete in time, although they all had their schema, as the original seed, in the mere self-development of reason” (CPR A 835 / B 863). In this passage, Kant is contrasting two different accounts of animal generation, namely a position obviously similar to that of metamorphosis, and a possibly preformationist one.

But the context of the passage is important here, because Kant is actually making a point about the structure of a cognitive system, not about its genesis. He is arguing that

---

<sup>7</sup> And it may be more than an analogy, if we are to believe Justin Smith: cf. 2.2.2.

it is impossible to regard systems as arising through the mere aggregation of knowledge or concepts, as is revealed by the following passage:

The scientific rational concept thus contains the end and the form of the whole that is congruent with it. The unity of the end to which all parts are related and in the idea of which they are also related to each other, allows the absence of any part to be noticed in our knowledge of the rest, and there can be no contingent addition or undetermined magnitude of perfection that does not have its boundaries determined a priori. The whole is therefore articulated (*articulatio*) and not heaped together (*coacervatio*); it can, to be sure, grow internally (*per intus susceptionem*) but not externally (*per appositionem*), like an animal body, whose growth does not add a limb but rather makes each limb stronger and fitter for its end without any alteration of proportion. (CPR: A 832-833/B 861)

Here, Kant is suggesting that systems develop through a process parallel to that of intussusception. We have encountered this position before in our discussion of Buffon's version of metamorphosis. Buffon postulated that in order to understand the organic process of nutrition, we need to postulate a kind of internally fashioning power such as intussusception, because mere addition of new parts could not explain the specificity of organic form. Buffon likely arrived at the notion of intussusception through his reading of Bourguet, who offered it in order to develop a preformationist theory of development (cf. Roger 1963: 376-377). The term itself, however, is substantially older, since it appears equally in the writings of the Jesuit Scholastics, possibly coined by Rodrigo de Arriaga: "plants and other animate things do not grow by juxtaposition: they grow *intus-sumptio*, that is, by attracting to themselves through their pores food divided into tiny particles, and converting it into themselves" (Arriaga; translated and quoted in Des Chene 2000: 61-62). The notion of intussusception therefore corresponds with the recognition that an organic whole has to integrate new parts in a different way than an aggregate, since it has to adapt these new forms to its own structure. This leads to the idea that in some way this structure is presupposed. We can therefore better understand Kant's comment on the metamorphosis theory here if we assume that he is alerting to the fact that even metamorphosists like Buffon have been forced to assume some pre-given structuring principle in order to allow for specifically organic structure and growth in the form of an interior mold. All that Kant needs to be read as saying, then, is that there need to be specific structural constraints on the growth of an organism on any account. But this can be in line with epigenesis as well, since it too holds that there are constraints, i.e. germs and dispositions, on the process of generation. For Kant, however, these constraints are determined by an idea of the whole. That the idea of the whole determines the process of generation is not a specifically preformationist idea either. Finally, the core idea here being that of intussusception suggests that here, like in the passage from the "Transcendental Analytic", Kant is mostly concerned with the idea

that the changes to an organic whole must be internally conditioned, and it is for this that the talk of dispositions was meant to allow.

On the whole then, I do not find that any of the adduced passages decide clearly in favor of the preformationist reading of Kant's philosophy, since they can also be read as expanding on points from his epigenesist theory. The extent to which one would hold the preformationist view as tenable depends therefore on what sense it can make of Kant's own position. I believe the reading of Kant's position as a mitigated preformationism is also flawed in the latter way, because it ends up in difficulties for two reasons. The first is that despite its insistence that Kant endorses a *mitigated* form of preformationism and of a priori-fixism, it fails to show how these forms are mitigated at all. The second is that it has to ascribe to Kant a consistent misinterpretation of his own views, or an inadvertency to the fact that they change, since he never uses the idea of preformation positively, and only ever uses the term epigenesis positively.

All the authors advancing or adopting the thesis want to stress that Kant departs from the Leibnizian interpretation of innateness in terms of innately present mental states, by advancing an idea of only latently present dispositions that tend to develop in one way or another, but are nonetheless somehow fixed and innate. The problem with this narrative is that it introduces a difference between Kant and Leibniz only by dramatically misconstruing Leibniz's theory. In fact, Leibniz's theory is better understood as committed to the existence of latent, innate conceptual capacities that only develop on occasion of experience and the reflection thereon. But this is precisely the account many have taken Kant to offer *as an alternative to Leibniz*.

The same goes for the issue of preformationism. The Sloan-Zammito thesis rests on the idea that Kant rejects a crude preformationism in favor of the refined version that can be found in the works of Albrecht von Haller and Charles Bonnet. As should be clear from my previous chapter, the idea of a crude preformationism is a retrograde construct, not based on the actual perception of the theory by 18<sup>th</sup> century theorists. Like in the case of Leibnizian innatism, many current readers ironically take Kant to be criticizing preformationism just to embrace what is in fact orthodox preformationism.

If this is true, then we have two major options open to us. The first is to admit that Kant's position is ultimately just one more version of Leibnizianism, just one more version of preformationism, and that Kant merely grossly misunderstood his own theory as a counterpart to such theories because he was a sloppy and/or malicious reader of his contemporaries. This conclusion may be very welcome to many philosophers now, since they have always felt difficulty in clarifying their opposition to transcendental philosophy's focus on the a priori and the categorial. Its major flaw is, however, that it implies that Kant could succeed in giving, in a work destined to German philosophical Academia, an account of Leibnizian philosophy that would immediately be recognized as flawed by the very members of German philosophical Academia. Moreover, it would consist in him offering an alternative to a philosophy by offering

that philosophy itself. Even for those willing to admit that great philosophers have their flaws, their oversights, even their stupidities, it should be outrageous to suggest that so big an oversight, so obvious a stupidity, could escape the attention of so many readers.

I suggest therefore that we take the second option by assuming that Kant was at least partly aware of the subtleties of Leibnizianism and preformationism, and wanted to develop a position differing from their refined versions. This allows us to entertain the idea that his commitment to historicity is real, even though the latter concept does not take the forms with which we are familiar from the early romantic and German Idealist paragons of historicism. It is perhaps by reading Kantian philosophy of biology through the Darwinian lens and Kantian philosophy of history through the Hegelian or Herderian lens that we have seen it as rigid. The consequence was probably that, in order to make sense of the fact that Kant believed himself to differ from Leibniz, we were forced to start substantially misreading the latter. In a peculiar twist of the historiography of philosophy, we have ended up reading the terms that Kant used in order to account for the developmental, the historical, the genetic aspect of his philosophy so that they ended up denying development, history and genesis. I suggest we postpone such harsh conclusions until we have examined the matter more fully.

### 3.3.2 The Epigenesis of Reason

For all its talk of unequivocal preformationist language in Kant's philosophy, the Sloan-Zammito reading is mostly concerned with explaining away the following sole unequivocal passage in the whole of Kant's epistemology, in which he explicitly asserts that his theory of knowledge is analogous to epigenesis, and refuses any account based on either metamorphosis or preformation. The passage is to be found in the second edition reworking of the Transcendental Deduction:

But this cognition, which is limited merely to objects of experience, is not on that account all borrowed from experience; rather, with regard to the pure intuitions as well as the pure concepts of the understanding, there are elements of cognition that are to be encountered in us a priori. Now there are only two ways in which a *necessary* agreement of experience with the concepts of its objects can be thought: either the experience makes these concepts possible or these concepts make 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 *selfthought* 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. (CPR B166-168).

This passage has attracted more attention than the allegedly preformationist one from the first edition, probably because the phrase “epigenesis of reason” is so suggestive, and because it may help us figure out what the specificity of Kant’s position is with regard to the alternatives he discerned. These alternatives, I submit, are either Lockean empiricism or Leibnizian rationalism. In offering the interpretation that what Kant is here calling epigenesis is in fact a form of preformationism, those upholding the Sloan-Zammito thesis are thereby arguing that Kant’s thesis is nearer to preformation than is usually thought. This raises several problems, some of which I have already raised.

First, it would have to be admitted that Kant moved from a preformationist position to one that he mistook to be epigenesist, but which was actually preformationist. But even if he mistook his later position to be epigenesist, this does not explain why he did not feel the need to change the passages from the first edition that are allegedly unambiguously preformationist. I think we can make better sense of these facts if we state that Kant was simply always epigenesist, and did not believe the passages from the first edition committed him to another position than epigenesis. I have shown, in the previous section, how he could have believed this.

Second, it is problematic to regard Kant as searching for some kind of mitigated, middle-way position in this passage in light of the fact that he does not offer his own view, that of the epigenesis of reason, as a middle ground between the preformationism-system of pure reason and the *generatio aequivoca* system. Instead, he presents Leibniz’ preformationist model of thought as the middle way between his own theory and that of Locke. Any reading of this passage has to clarify what Kant could have meant by this. I will argue that the crucial idea here is one which already appeared, although confusedly, in the first edition, namely that of the categories being “self-thought”. In this way, we can also explain why Kant took this passage to be a clarification of the first edition comments, rather than a retraction; a reading which explains why they appear alongside each other in the second edition. I will show first how Leibnizian rationalism and preformationism could be regarded as parallel, then how Kant could conceivably make a parallel between Locke’s empiricism and metamorphosis and finally what this implies for his analogy between transcendental idealism and epigenesis.

The first thing required to make the parallel between the two epistemological theories of empiricism and rationalism match up with the alternative between metamorphosis and preformation is to indicate the origin of Kant's insistence on the empiricist-rationalist divide. If it is obvious to many contemporary readers that Kant is entitled to present the early modern debate on theory of knowledge as shaped by this fundamental opposition, then this is the result of almost two centuries of historiography based on the histories of philosophy proposed by Kantian and post-Kantian German philosophers. In the course of the late 20<sup>th</sup> century, however, this opposition has lost its obviousness to historians of philosophy. I submit that this is because we have all too rarely identified the specific textual basis Kant had for proposing the opposition, namely Leibniz's *New Essays Concerning Human Understanding*. In this work, Leibniz intends to give a section-by-section, chapter-by-chapter response to Locke's original landmark *Essay Concerning Human Understanding*. For us, it is interesting to note that although Leibniz himself wrote this book in 1704, it did not appear in print for the larger public until 1765. I will here briefly argue for the fact that it was this work, not the works of David Hume, that awoke Kant from his dogmatic slumber. This appears first of all from the passage from the prolegomena where Kant starts to discuss the idea of a dogmatic slumber:

Since the Essays of Locke and Leibniz, or rather since the rise of metaphysics as far as the history of it reaches, no event has occurred that could have been more decisive with respect to the fate of this science than the attack made upon it by David Hume. (AA IV: 257)

In this passage, Kant is indicating two major landmarks in his philosophical landscape. The first landmark is made up by the Essays of Locke and Leibniz. The reason why Kant can take these together is because the two works are intertwined, the latter being a commentary on and response to the former. The other landmark is the work of David Hume. The problem is now that there is a problem of temporality occurring here, since Leibniz's Essays appeared only in 1765, around the time when we tend to situate Kant's awakening from dogmatic slumber. Is Kant then misrepresenting the history of philosophy in the most basic way, namely by getting the *dates* wrong? I believe this would be a mistaken conclusion, since Kant notes the temporal anomaly in the very passage where he introduces the idea of a dogmatic slumber, saying that: "I freely admit that the remembrance of *David Hume* was the very thing that many years ago first interrupted my dogmatic slumber and gave a completely different direction to my researches in the field of speculative philosophy". (AA IV: 260) Mind the word that is usually fatefully omitted from references to this passage, namely that of *remembrance*. Kant is *not* saying that he was first awakened by reading David Hume, but that it was the *memory* of David Hume's philosophy at a later moment in his career that awakened him. I propose the following reading of this passage.



When Leibniz's *New Essays* finally appeared in 1765, it entered a philosophical landscape where the differences between Leibniz's and Locke's approaches to philosophy were recognized as supremely important. The two philosophers represented two different ways of doing Enlightenment philosophy. In many specific debates, the Lockean and the Leibnizian options found themselves to be direct opposites. The *New Essays* served to make sense of these oppositions, and offered glosses on how exactly Leibniz saw them. When Kant read the book shortly after it appeared (and there is every reason to assume that he did), he found it enormously helpful to understand and systematize the many debates and tensions in 18<sup>th</sup> century philosophy.

The most specific feature of the *New Essays*, however, is that it contains a much more elaborate and refined picture of Leibniz's epistemology. Until then, Leibniz was read more through the lens of his metaphysics, but here, in his dialogue with the more metaphysically reticent Locke, the debate focused on more recognizably epistemological issues. Part of the task Leibniz seems to have set himself in this book is to show how Locke's theory of knowledge makes it inconceivable how we can ever come to acquire knowledge through the mechanisms of association and reflection alone. For this reason, he suggests that the admittance of innate ideas and dispositions are indispensable, even to Locke. In a lecture on Logic, Kant is explicit in assigning a central place to the Locke-Leibniz debate in the history of logic and metaphysics:

Locke became famous through his *Essay Concerning Human Understanding*— he speaks there of the origin of concepts, <but> this really does not belong to logic, but rather to metaphysics. The result of his investigations: Everything derives from experience. But it does not follow at all from this that concepts can be *displayed* only in experience. - Then Leibniz entered the picture. Although actually having written no logic, he nonetheless did much to illuminate concepts (he wrote in defense of his countrymen against the Englishman Locke). In his works he expressed ideas which subsequently moved Wolff to his system. (AA XXIV: 701)

The theme of the possibility of knowledge is recognizable as the major turning point in Kant's career. It also explains how Leibniz's discussion would cast light, in retrospect, upon Hume's philosophy, for, if Leibniz is right that Locke's theory of ideas makes it impossible to explain knowledge, then David Hume can be read as drawing this radical conclusion from it and embracing the resulting skeptical view. Kant's interest in Hume is therefore likely to have been reignited by his reading of the *New Essays*, which explains why he speaks of the *remembrance* of Hume as having awakened him from his dogmatic slumber.

This account yields the idea that, in reading Kant's comments on the course of the history of philosophy or the lay of the philosophical landscape in his own century, we should keep in mind the *New Essays*, for it is likely this book that has shaped Kant's critical turn more than any other. Further on in this dissertation, we will find other

examples of the influence of the *New Essays*, but for now we will focus on the opposition between the Lockean and the Leibnizian theories of knowledge.

The parallel between Leibniz's epistemology and the preformationist theory of generation is easily made, since Leibniz himself made it in the following manner: "it [i.e. understanding] is not a bare faculty, consisting in a mere possibility of understanding those truths [i.e. eternal and necessary truths]: it is rather a disposition, an aptitude, a preformation, which determines our soul and brings it about that they are derivable from it". (G V: 77; Leibniz 1996: 80) This of course lends further support to the idea that Kant is thinking, in the *Critique of Pure Reason*, primarily of the *New Essays* as the source for his interpretation of his predecessors in epistemology. Given this passage and its context, Kant is entitled to believe that Leibniz sees the parallel as follows.

Locke argues that no innate ideas are required in order to account for knowledge, and that only the bare faculty of knowledge is required in order to extract from a manifold of empirical sensations all the concepts which we can claim to have. Leibniz, however, insists that no (finite) amount of empirical information can yield that indispensable element of knowledge that is necessity. Concretely, the necessary principles that structure and guide knowledge cannot be derived from mere experience, and must therefore have a different origin if knowledge is to be possible at all. That is why Leibniz suggests that we have been furnished with them at the moment of creation, and that although we only come to explicitate them on the occasion of experience, they do not derive from this experience, but from the internal development of knowledge. Similarly, Leibniz wishes to argue that the specific structure of living systems cannot be brought forth by the general principles of mechanics alone (or at least by any finite series of mechanical operations), which is why he postulates that all organic structures are individually preformed at the moment of creation, and that although these structures develop on the occasion of fertilization and nutrition in a sufficiently similar individual, they are not produced by these processes.

It is therefore Leibniz, and not Kant, who originally, though tacitly, suggested the parallel between Lockean empiricism and metamorphosis. Remember that for Harvey, the system of metamorphosis was associated with that of spontaneous generation, because in denying the specificity of the organic constraints on development it is committed to the idea that, in principle, life might arise out of (appropriately concocted or prepared) unliving matter. The view equally led Buffon to allow spontaneous generation through his association with Needham. Like metamorphosis, empiricism believes that we can get an organized body (in this case of knowledge) simply from the given materials (the empirically given) and the simple operations at work in them, like association. The idea that Locke saw a parallel between his austere set of ideas and mental operations and the mechanist ideal of a general kind of matter governed by a limited set of mechanical principles, could have been suggested by the following passage from the *Essay Concerning Human Understanding*:

These *simple* ideas, the Materials of all our Knowledge, are suggested and furnished to the Mind, only by those two ways above mentioned, viz. *Sensation* and *Reflection*. When the Understanding is once stored with these simple *Ideas*, it has the Power to repeat, compare, and unite them even to an almost infinite Variety, and so can make at pleasure new complex *Ideas*. But it is not in the Power of the most exalted Wit, or enlarged Understanding, by any quickness or variety of Thought, to *invent or frame one new simple Idea* in the mind, not taken in by the ways before mentioned: nor can any force of the Understanding, *destroy* those that are there. The Dominion of Man, in this little World of his own Understanding, being much what the same, as it is in the great World of visible things; wherein his Power, however managed by Art and Skill, reaches no farther, than to compound and divide the Materials, that are made to his Hand; but can do nothing towards the making the least Particle of new Matter, or destroying one Atome of what is already in Being. (Locke 1975: 119-120)

The idea that this picture of knowledge fails to give us the element of necessity required to speak of knowledge in any important sense of the word, would be constantly voiced, but of course reached its peak when Hume suggested that the Lockean theory could indeed not offer any strong connection between ideas:

As all simple ideas may be separated by the imagination, and may be united again in what form it pleases, nothing wou'd be more unaccountable than the operations of that faculty, were it not guided by some universal principles, which render it, in some measure, uniform with itself in all times and places. Were ideas entirely loose and unconnected, chance alone wou'd join them; and 'tis impossible the same simple ideas shou'd fall regularly into complex ones (as they commonly do) without some bond of union among them, some associating quality, by which one idea naturally introduces another. (Hume 2009: 12)

Hume objects to Locke's atomism of the mind because it makes the manner and regularity of the association of ideas incomprehensible. This regularity suggests that ideas bond more readily with certain ideas than with certain others. Hume submits here that we need a principle that can account for the cohesion of ideas and the apparent selectivity of that cohesion. He next introduces precisely such a principle, stating that it is "a kind of *attraction*, which in the mental world will be found to have as extraordinary effects as in the natural, and to show itself in as many and various forms" (Hume 2009: 14). A major innovation is that he advanced *custom* as the principle for the stronger connections between ideas that are otherwise mere matter of fact, and that he characterized custom as an analogue to the force of attraction in Newtonian Mechanics. This would have reinforced the parallel for readers like Kant, since both Hume and theories of metamorphosis invoke a force analogous to Newtonian attraction to explain that for which the mechanical picture could not account otherwise. Just like Buffon and Maupertuis postulated further principles after the example of gravitation in order to

explain the emergence of the intricate structure of an organism, so Hume postulated a further principle in order to explain the emergence of the kinds of connections we consider essential to knowledge.

It is clear that Kant shared with Leibniz a broad skepticism with regard to the potential of Lockean empiricism to account for knowledge. For him, this skepticism was enhanced by the fact that Hume seemed to believe, and was widely held to be entitled to believe, that this failure should lead the empiricist to embrace skepticism about the possibility of knowledge. In §13 of the first *Critique*, Kant explicates his criticism of Locke in the following manner:

in the case of these concepts, as in the case of all cognition, we can search in experience, if not for the principles of their possibility, then for the occasional causes of their generation, where the impressions of the senses provide the first occasion for opening the entire power of cognition to them and for bringing about experience, which contains two very heterogeneous elements, namely a matter for cognition from the senses and a certain form for ordering it from the inner source of pure intuiting and thinking, which, on the occasion of the former, are first brought into use and bring forth concepts. Such a tracing of the first endeavors of our power of cognition to ascend from individual perceptions to general concepts is without doubt of great utility, and the famous Locke is to be thanked for having first opened the way for this. Yet a deduction of the pure *a priori* concepts can never be achieved in this way; it does not lie down this path at all, for in regard to their future use, which should be entirely independent of experience, an entirely different birth certificate than that of an ancestry from experiences must be produced. I will therefore call this attempted physiological derivation, which cannot properly be called a deduction at all because it concerns a *quaestio facti*, the explanation of the possession of a pure cognition. It is therefore clear that only a transcendental and never an empirical deduction of them can be given, and that in regard to pure *a priori* concepts empirical deductions are nothing but idle attempts, which can occupy only those who have not grasped the entirely distinctive nature of these cognitions. (CPR A 86-87 / B 118-119)

Kant rejects Locke's empiricist approach to knowledge because it fails to *justify* the concepts it *explains*. Mind that the point here is not that such an account is worthless, nor that it cannot work for some concepts. It is very useful for showing how we come to acquire most of our concepts, but it cannot account for those concepts which themselves form the necessary capacities for the acquisition of concepts. On Kant's account, all our empirical concepts are of course empirically derived, but they do not derive their status of concepts, the role they get to play in knowledge claims, from their empirical sources. It is the categories that are responsible for that particular aspect of concept acquisition. And if this is true, then of course these categories cannot themselves be empirically acquired, on pain of infinite regress.

The prima facie similarity between the diagnoses of the failures of Lockean empiricism offered by Leibniz and Kant has undoubtedly furthered the idea that the latter embraced some form of the theory espoused by the former. The distinction is then supposedly that, according to Leibniz, *all* concepts are innate, whereas for Kant only those necessary for concept acquisition are. But there are problems with this reading, for it would suggest that Kant took an intermediate position between Locke and Leibniz, claiming that some concepts are somehow innate or preformed whereas others are not. This fails to explain why Kant describes preformationism as a middle way between his own position and that of Locke. In fact, Kant cites an entirely different reason for his dissatisfaction with the Leibnizian preformation theory of concepts. First of all, he only talks here about those concepts necessary for knowledge, in his terminology the categories, and charges Locke with failing to indicate how we would come about them. The idea seems to be that, in absence of any prior concepts regulating concept formation and acquisition, all concepts would remain merely subjective, and lack the objectivity they need. If we were therefore to acquire the concepts that correspond, in the Lockean theory, with the categories, these would themselves be merely subjective. Lockean empiricism fails, in Kant's view, to account for the *objectivity* of the categories.

Whereas the criticism of Locke is fairly easy to identify, given that Kant expands on it elsewhere and given its connection to the challenge of Humean skepticism, the criticism of Leibniz is more easily misunderstood. Kant offers several objections, but first it is important to note that he takes the categories to be, on the preformationist account, "subjective predispositions for thinking". In identifying the Leibnizian version of the category as subjective, he is stating that the preformation theory fails to solve the problem of the origin of objectivity. I submit that this is so because the bruteness of their acquisition is no less reduced. According to Kant, Leibnizian innate ideas of necessity have not been arrived at in any way, but are merely factual givens of our human or individual nature. Their objectivity (in *casu* their correspondence with reality) would be entirely accidental to them, for they would not actually relate to the objects about which they are, and, as a result, "in such a case the categories would lack the necessity that is essential to their concept". Indeed, Leibniz's preformationism suggests that our ideas and concepts of things have no real relation to those things: they correspond with them only because God has implanted in us certain ideas and made sure that they correspond with their counterparts independent of all interaction. Without such a divine warrant, we never have knowledge at all, as our concepts lack justification.

Kant adds to this another objection, that on this account "no end can be seen to how far one might drive the presupposition of predetermined predispositions for future judgments". The idea behind this comment is arguably that, once the concepts necessary for concept acquisition have already been admitted as innate and preformed,

there is no reason not to expand this to all empirical concepts, as Leibniz in fact did. This suggests that Kant identified the preformation theory with the theory that the categories are preformed, not with the theory that all concepts are preformed, and charged the former theory with its tendency to slip into the latter. It is hard to maintain the interpretation according to which opposed to the preformationist theory his own theory that the categories are be preformed, whereas empirical concepts are acquired through the employment of the categories, since the latter theory is precisely what Kant calls preformationism in this passage.

But then what does Kant mean when is speaking of the epigenesis of reason? As A. C Genova (1974) has suggested, Kant seems determined to warrant the *autonomy of reason* in its yielding of the categories. Such an autonomy would not be helped by replacing the external imposition on the part of experience to the equally external imposition on the part of God or any predetermining factor. This is supported by the fact that Kant contrasts the Leibnizian “subjective predispositions for thinking” with his categories as “*selfthought* a priori first principles of our cognition” (Kant’s stress). The importance of the epigenesist story is, then, for Kant, that the categories are not given to us by experience, nor originally given, but are due to reason itself. Secondly, this is perfectly in line with the passages we discussed from the first Critique, all of which could be read as stressing the internal origin of the categories rather than the preformation of the categories.

Finally, it allows us to make sense of Kant’s utterance that preformation is a middle way between epigenesis and metamorphosis. Epigenesis asserts that the emergence of the organism is due to the specifically organic constraints, and so due to the organism itself, which, through some form of circular causation, determines its process of genesis. Metamorphosis instead suggests that the level of the organic is superfluous in explaining the emergence of organic structure, since the materials and the principles governing them suffice. Preformation agrees with metamorphosis that the level of the organic is not active in the process of generation, and with epigenesis that the organic structure cannot be accounted for by the materials and their principles. Hence, it suggests that organic form is externally imposed on the matter it structures. In the same way, the preformationist account of knowledge does not seek the origin of knowledge in the mere materials of knowledge, nor in the form of knowledge itself, but claims that the form of knowledge is externally imposed.

We have here argued that Kant’s epigenesis of reason is meant to be a position that greatly differs from Leibniz preformationist theory of knowledge, and that one major aspect of difference is that for Kant, the categories must be somehow due to the internal structure or force of reason, not due to some external imposition by nature or by God. It will only be in the final chapter of this dissertation that we will be able to flesh out this theory of epigenesis. Before concluding this first part, however, I want to further articulate our results so far by relating transcendental idealism to the idea of historicity.

### 3.4 Transcendental Idealism and Historicity

In this section, I want to briefly indicate how the idea of history arising through our analysis of Kant's philosophy of history and the idea of epigenesis as it appears in the transcendental deduction can relate. I will first set the stage for my comments by outlining some of Jacques Derrida's insights into the relation between transcendental idealism and historicity. Then, I will relate the biological theory of epigenesis, the philosophy of history and the idea of the epigenesis of reason on the basis of the four relevant aspects of Kantian epigenesis identified in the previous chapter. This discussion is not meant to be decisive, nor fully enlightening, but it serves to indicate what challenges and questions lie ahead of us in the next four chapters.

The central question in this dissertation is whether we can take Kant's epigenesis of reason as suggesting a theory that allows for the historicity of reason without giving up knowledge's claims to objectivity. Such a theory is very difficult to offer, but, as Jacques Derrida has said in the context of his introduction to Edmund Husserl's *Origin of Geometry*, "If we take for granted the philosophical nonsense of a purely empirical history and the impotence of an ahistorical rationalism, then we realize the seriousness of what is at stake" (Derrida 1989: 51).

With this last quotation, we have found an entry point into the idea of transcendental history. After all, it was Derrida's concern to reveal how, in Husserl's early work, the transcendental could be harmonized with the historical. The main move in this idea is, of course, that only the transcendental can allow for the historical:

If the history of geometry were only the development of a purpose wholly present from the beginning, we would have to deal only with an explication or a quasi-creation. We would have on one side a synchronic or timeless *uchronique* ground and, on the other side, a purely empirical diachrony with its indicative function but without any proper unity of its own. Neither pure diachrony nor pure synchrony make a history. The rejected hypothesis is once more that of a complicity between "Platonism" and empiricism. (Derrida 1989: 62)

The importance of this passage for my discussion is of course immediately obvious, since the position described here as "Platonism" is precisely the Leibnizian position, which tends to deny that concepts have a history, since they are atemporally preformed and therefore come to be instantiated only through a quasi-history, a merely *pro forma* settling of a fully predetermined matter.<sup>8</sup> Interestingly, however, Derrida goes on to

---

<sup>8</sup> At a certain point, Derrida suggests that Husserl advances on Kant through allowing for the historical, whereas Kant is still caught up in the Platonist picture (Derrida 1989: 39-41). My discussion in this chapter is

deny that empiricism can account for history, since it cannot show how the sequence of empirical occurrences and experiences can constitute a unity such that it can yield specific conceptual results. Later in the essay, Derrida announces that history is made such a unity, is made historical, by the Idea, which he himself insists should be taken in the Kantian sense. This forces him to respond to the obvious criticism that history is now made one by something that is itself ahistorical: the idea, with the following pellucid comment:

Undoubtedly the Idea and the Reason hidden in history and in man as “animal rationale” are eternal. Husserl often says this. But this eternity is only a historicity. It is the possibility of history itself. Its supratemporality-compared with empirical temporality-is only an omnitemporality. The Idea, like Reason, is nothing outside the history in which it displays itself, i.e., in which (in one and the same movement) it discloses and lets itself be threatened. (Derrida 1989: 142)

Derrida’s analysis of the role of the idea in Husserl suggests that the idea, which is supposed to be untemporal, is not, after all, unhistorical, and is in fact open to contingencies. I believe a similar situation holds for Kant’s philosophy, since such a situation makes the best sense of what are otherwise overt contradictions in his thought. This is brought out well by Derrida’s reference to the idea of reason in history. As I have argued in section 3.2., Kant’s philosophy of history presents reason not as some fixed, eternal faculty, but as the faculty for historicity. Kant argued that we can never know our predispositions, but that we can know reason. Reason, here, is the faculty that causes us to leave our initial stage and the fixed confines of our faculties and develop ourselves and our potentials through a historical process. Derrida correctly saw that for transcendental idealism, reason is “the possibility of history itself”. Reason should therefore not be regarded as a preformed faculty, one that is granted by the atemporal, but rather as one that inhabits our history as a virtuality, much like the germs and predispositions, as internal constraints on the generative faculty, inhabit it as virtually preformed elements.

Secondly, in this analysis, the idea of reason is the “possibility of history itself” because it is capable of establishing a continuity in a historical development which ties them together as different stages in one history, and not just different disjoint times. Such continuity is arguably only possible if there are some constraints on how two subsequent phases may relate and may differ to make them importantly different forms of the same. In Kant’s philosophy of history, the germs and predispositions are intended

---

intended of course to counter such a reading, and reveal that what is supposed to be original in Husserl’s version of the transcendental is already present in Kant’s, and is therefore a fundamental tenet of transcendental philosophy as such.



to provide precisely not only the fact that we tend to historical development, but also that this historical development is tied together as a continuity.

Reason is not, of course, fully historicized by these two glosses, these two shifts between the “Platonic” and the “Transcendental” picture. In order to be genuinely historical, reason must be open to contingency, which means that it must be amenable to change itself. An important feature of Derrida’s reading is precisely that, in history, novelty can arise. In order for this to be possible, reason cannot have a rigorous structure such as it is supposed to have on Platonic accounts: its omnitemporality means that it, although it is not limited in time, is nonetheless open to the effects of time, of history. A similar stress is present in Kant’s theory of predispositions, because it entails that our very faculties can change through a historical process, and that it is not predetermined to which constraints this change is subjected, much like the theory of predispositions is supposed to allow for the plasticity of development and of species in Kant’s theory of epigenesis. According to Kant, the very ideas of education and of culture suggest such malleability of our faculties.

Such openness to contingency, however, should not be misconstrued as direct openness to external influence. In the quoted passage, Derrida is careful to insist that it is reason itself which “lets itself be threatened”. This means that the openness to contingency is internal to reason, and that contingency arises in and through the very same process whereby reason develops itself. Derrida undoubtedly stipulates this because the internality is a condition for reason taking this contingency into account: as long as the contingency remains external, it can be dismissed as mere historical accident and anecdote. Thus, the place and function of contingency must lie in the heart of reason itself. A similar idea can be found in Kant’s theory of epigenesis, where it is the generative force itself which bears, within itself, and through its workings, the possibility of change. The theory of predispositions, as we saw in the previous chapter, entails that contingency should be viewed as merely occasioning the internal capacity for change, rather than bringing it about directly. It thus turns out that the criteria Derrida identified as prerequisites for a transcendental philosophy that can accommodate historicity, are importantly analogous to the main tenets of Kant’s theory of epigenesis as identified in the previous chapter.

### 3.5 Conclusion

In chapter, I have dealt with the role of Kant’s talk of epigenesis in his philosophy of history and his epistemology, and shown that the four traits of Kant’s theory of epigenesis identified in chapter 3 allow us to understand this connection. This has set us

on the path for a better appreciation of Kant's theory of the epigenesis of reason as a transcendental epistemology that is open to historicity. Nonetheless, our analyses here do not in themselves yield such a picture, for it remains to be seen *how* Kant's reason can be epigenetical in such a manner. In order to answer this question, we will have to turn our attention to the nature of the systems that Kant would have to hold as epigenetical, which is the topic of the next part of this dissertation.

**Part II:**

**Organics, Systematicity and Purposiveness**



## Chapter 4 The Structure of the Organic: Natural Purposiveness and Circularity

*The fact that complex systems can be approximated (albeit locally and temporarily) by simple ones is a crucial one. It explains precisely why the Newtonian paradigm has been so successful, and why, to this day, it represents the only effective procedure for dealing with system behavior. But in general, we can also see that it can supply only approximations in general, and in the universe of complex systems, it amounts to replacing a complex system with a simple subsystem.*

- Robert Rosen

In the previous three chapters, I have argued that Kant's notion of epigenesis was intended to yield an account of why the transcendental, the structural, can be historical in the sense of being amenable to change and to contingency. It is now time to turn our attention to these dynamic structures. I will begin elaborating Kant's account of dynamic structure with an analysis of Kant's paradigm instance of such a structure, namely the organism. The concept of an organism, we will see, is peculiar in Kant, because it is intimately linked to that of a natural purpose, although it would be mistaken to identify the two. To show what is at stake in Kant's theory of natural purposiveness, I will begin, in Section 4.1, with a necessarily all too brief historical overview meant to indicate the issues surrounding the notion of teleology in Kant's time. In subsection 4.1.1, I will argue that the difference between Platonic and Aristotelian teleology is that the former always involves intentionality, whereas the latter can be non-intentional and yet non-mechanical. The theistic tradition inherited Plato's concept and rejected Aristotle's, thereby suggesting that all natural teleology derives from a transcendent artificer, i.e. God. In subsection 4.1.2, I will argue that mechanism's critique of teleology was based on, rather than directed against, the theistic position. In section 4.2, I will argue that Kant's concept of the "natural" in

“natural purpose” refers primarily to the non-intentional nature of the purposiveness involved. I will also argue that, once we understand this, we will need to let go of the idea that the mereological relation exhibited by natural purposes is that of the primacy of the whole of the parts, but rather accept that in natural purposes no such priorities can be assigned. In subsection 4.2.1, I will show how Kant delineates the kind of purposiveness with which he is concerned. I will then discuss, in subsection 4.2.2, Kant’s preliminary characterization of the concept of a natural purpose through his example of a tree. This discussion will reveal that Kant thinks of natural purposes along Aristotelian lines, in that they are entities which need to be judged as somehow bound or obliged to maintain themselves as specific individuals of specific kinds, although this specificity is subject to some plasticity. In subsection 4.2.3, I will discuss Kant’s determination of the concept of natural purpose proper, showing that it cannot simply be understood through mechanism alone, nor through an analogy with art or intentional design. Natural purposes exhibit a circular causality which obtains between the multiplicity of their parts, and which results in the mutual (self)-specification of these parts. Since such a concept is incomprehensible to man, having no analogue with any of the two forms of causality with which we are familiar, namely mechanical causation and intentional action, there is a problem of how we are supposed to put it to use. This problem manifests itself in the antinomy of teleological judgment, which I discuss in section 4.3. In 4.3.1, I discuss the initial presentation of the antinomy in §§70-71 of the *Critique of the Power of Judgment*, and argue that it does not yet furnish a solution to the antinomy. The problem that remains is that, although the antinomy does not lead to a metaphysical contradiction, it does result in a conflict between two non-optional maxims for our inquiry into nature. In 4.3.2, I present §§72-74 as intended to explore possible solutions to the second challenge of the antinomy, and show why Kant ultimately found the assumption of design to be most helpful. In 4.3.3, I argue that Kant’s solution to the antinomy is the following: although natural purposes suggest a different type of causality than those familiar to us, and although this type of causality is not strictly impossible, we cannot comprehend anything in this manner. That is why we adopt the language of design in our inquiries: not because all teleology is intentional but rather because we can only *mechanically* reason about *teleology* if we consider it as intentional. I also discuss Kant’s reason why we cannot think of teleology otherwise than in this form, namely his conviction that humans are irreducibly divided between sensibility and understanding. I will conclude by noting some features of Kant’s concept of natural purpose that are relevant to the next chapters.

## 4.1 From Natural Teleology to Natural Theology

It has become an enduring tendency of philosophers to link Aristotle's philosophy, and his philosophy of nature in particular, to the concept of teleology. Like many enduring tendencies, however, they have become ingrained through repetition rather than through understanding, thereby impeding our understanding of what went on in the turn from the Medieval to the Early Modern conception of natural philosophy. An important aspect of this turn, we are told from early on in our philosophical development, is the banishment of final causation as an irreducible explanatory category. For Aristotle, we know, final causation is one of the four basic causes, one which is irreducible to the others, and indispensable for any proper explanation. Mechanical Philosophy, on the other hand, denies that indication of the final cause is indispensable and/or irreducible. It is also often believed that Aristotle's invocation of final causation is mostly motivated by concerns that we now regard as biological. The Aristotelian conception of the soul is then regarded as the notion that explains the agent behind the teleology purportedly exhibited by organisms.

In this section, I would like to discuss the ways in which this picture is right and the ways in which it is wrong. In the first subsection, I will argue that whereas Aristotle believed teleology to be potentially unintentional, Medieval Aristotelianism followed Plato in insisting that teleology always involves intentionality. As a result, it became increasingly difficult to conceive of final causes at work in processes where no intentionality seemed to be involved, except by referring to divine intentions concerning natural processes. This evolution led to the assumption that processes that do not involve a soul are only indirectly teleological, whereas those that do involve a soul can be directly teleological by referring to the intentions formed in that soul. I will argue that this led to the problem of the precise function of the vegetative soul.

In the second subsection, I will argue that mechanist philosophers relied on the scholastic transformation of teleology in order to eliminate natural teleology altogether or defer it completely to considerations of divine intentions. I will focus specifically on the contrasting accounts of final causation in Descartes and Boyle. I will argue that Boyle's account was the most important doctrine on final causation from the late 17<sup>th</sup> to the early 19<sup>th</sup> century because of its continuation in the programme of natural theology of the Boyle lectures. This is important because I believe Kant intended his account of teleology as an alternative to the particular wedding of teleology and theology in the natural theology of the Enlightenment.

### 4.1.1 The Naturality of Teleology

Aristotle's focus on teleology and finality in natural philosophy was not a complete invention of his, for here as in many other debates, he was influenced by Plato, who had already voiced, albeit through his mouthpiece Socrates<sup>1</sup>, objections against the materialist explanations offered by the presocratics. In the *Phaedo*, Socrates divulges to Cebes his one-time interest in the philosophy of Anaxagoras, for the following reason:

I once heard someone reading from a book, as he said, by Anaxagoras, and asserting that it is mind that produces order and is the cause of everything. This explanation pleased me. Somehow it seemed right that mind should be the cause of everything, and I reflected that if this is so, mind in producing order sets everything in order and arranges each individual thing in the way that is best for it. Therefore if anyone wished to discover the reason why any given thing came or ceased or continued to be, he must find out how it was best for that thing to be, or to act or be acted upon in any other way. On this view there was only one thing for a man to consider, with regard both to himself and to everything else, namely the best and highest good, since both were covered by the same knowledge. (97b-d)

Socrates was thus mostly enthusiastic about the fact that Anaxagoras seemed to take recourse to teleology in his explanation by allowing for the fact that mind (*nous*) was the primary cause of everything. This form of teleological explanation requires that one explains the structure of a thing or the processes which occur in it in terms of how they serve that thing, how they envisage the "good" of the thing. But this enthusiasm, Socrates continues, was soon dispelled:

As I read on I discovered that the fellow made no use of mind and assigned to it no causality for the order of the world, but adduced causes like air and aether and water and many other absurdities. It seemed to me that he was just about as inconsistent as if someone were to say, the cause of everything that Socrates does is mind – and then, in trying to account for my several actions, said first that the reason why I am lying here now is that my body is composed of bones and sinews [...] But to call things like that causes is too absurd. If it were said that without

---

<sup>1</sup> In this section, I will ignore the interpretative issues surrounding the fact that Plato usually advanced ideas or arguments through the mouth of characters engaged in a narratively complex dialogue. This fact, indeed, requires us to be careful in assessing the dialectical role Plato himself intended for the assertion or argument. I believe I can relatively safely bracket these issues here because my concern is more with the influence of the Platonic Corpus than with the internal logic of Plato's theory. The passages I will highlight have often been taken as Plato's own considered opinion by the authors with whom I will be engaging in the course of this section.



such bones and sinews and all the rest of them I should not be able to do what I think is right, it would be true. But to say that it is because of them that I do what I am doing, and not through choice of what is best – although my actions are controlled by mind – would be a very lax and inaccurate form of expression. Fancy being unable to distinguish between the cause of a thing and the condition without which it could not be the cause! It is this latter, as it seems to me, that most people, groping in the dark, call a cause – attaching to it a name to which it has no right. (98b-99b)

What Socrates finds objectionable in the presocratics is their omission of a proper explanation for the order exhibited by the processes they intend to describe. In other words, they fail to show how mind is operative in nature, and therefore fail to identify the proper cause at all.<sup>2</sup> What is needed according to Socrates (and arguably Plato) is an account of natural processes that shows how the material processes are employed by mind in its activity.

Plato himself would propose such an alternative style of explanation in natural philosophy in his *Timaeus* (cf. Wright 2000: 10-11), which distinguishes between a primary and a secondary sense of cause:

All these [material causes] are to be reckoned among the second and co-operative causes which God, carrying into execution the idea of the best as far as possible, uses as his ministers. They are thought by most men not to be the second, but the prime causes of all things, because they freeze and heat, and contract and dilate, and the like. But they are not so, for they are incapable of reason or intellect; the only being which can properly have mind is the invisible soul, whereas fire and water, and earth and air, are all of them visible bodies. The lover of intellect and knowledge ought to explore causes of intelligent nature first of all, and, secondly, of those things which, being moved by others, are compelled to move others. And this is what we must do. Both kinds of causes should be acknowledged by us, but a

---

<sup>2</sup> It is interesting to note that Socrates is not claiming that the presocratics fail to account for or explain Mind. Mind itself is not an explanandum here, but an explanans: what they fail to show is how Mind is or gets to be explanatory in nature. This brings up the question as to what exactly the explanandum, or the aspect of the explanandum, for which they would thereby fail to account, could be. An initial response would be that order is that explanans. I believe this is true, but that it is too brief an answer. What Socrates himself advances as the explanandum in the quoted passage is “doing what I think is right”. I submit that it is the notions of right, of good, and the normative dimensions they employ, that stand in need of explanation. For Socrates, a proper explanation of his sitting there in prison would require an explanation of the juridical context of his trial as described in the *Apology*, and the reasons against avoiding execution he adduces in the *Crito*. I also believe a further dimension needs to be taken into account, namely why it is also *best for Socrates* to be there in prison. This is in fact the primary topic of the *Phaedo*: the idea that Socrates is not harmed by his execution, because of the doctrine of the immortality of the soul. In this way, the concept of “right” in these passages from the *Phaedo* (re)collects all the dimensions of normativity for which mind is supposed to account.

distinction should be made between those which are endowed with mind and are the workers of things fair and good, and those which are deprived of intelligence and always produce chance effects without order or design. (46c-e)

In this passage, *Timaeus* (arguably Plato's mouthpiece in this dialogue), distinguishes between the primary causes of things, which account for their goodness and fairness, and which are due to the action of mind (*nous*), and the secondary causes, which describe the material processes at work and are due to the action of necessity (*anankè*). He stresses that a proper explanation requires both causes to be given, but that the relation between the causes must be such that the secondary causes are subservient to the primary causes. This distinction therefore coincides, at least roughly, with that from the *Phaedo* "between the cause of a thing and the condition without which it could not be the cause".

Plato's account of teleological explanation is fateful both for its account of the explanans and for its account of the explanandum. The explanans of teleology, for Plato, is mind or intelligence. The Myth of the *Timaeus* suggests that we take this quite literally<sup>3</sup>, since it refers to the intentional actions of a transcendent creator of the universe, the so-called Demiurge, who fashions the visible universe as an icon of an intelligible (and, for Plato, therefore invisible) paradigm.

The explanandum is the beneficial order exhibited by the cosmos and its constituents. This order is explainable only because "God desired that all things should be good and nothing bad, so far as this was attainable", and considered that "[order] was in every way better than [disorder]" (30a). The cosmos as a whole, *Timaeus* suggests, reveals the hand of the Demiurge for two major reasons, namely its closure and its proportionality. The closure is due to the fact "the creator conceived that a being which was self-sufficient would be far more excellent than one which lacked anything" (33d). The proportionality of the universe, on the other hand, reveals that its cause had an advanced knowledge of mathematics. More interesting for our present discussion is, however, why Plato believes that the explanation in terms of Mind is required in the case of living beings. The *Timaeus* contains lengthy discussions of the anatomy and physiology of living things, offering possible or probable secondary causes employed by the Demiurge in fashioning them, and reasons for his choice of these secondary causes. Thus, he seems to believe that the intelligence of the Demiurge is primarily responsible for the structures of these animals, not the workings of these animals.

Sarah Broadie has developed an interesting interpretation of Plato as holding that the intelligence of the Demiurge only operated at the creation of the universe, but that he

---

<sup>3</sup> This of course does not itself answer the question of whether or to what extent the Myth itself need to be taken literally.

withdrew upon the completion of his creative act in order to leave nature to its ordinary course:

[The cosmic animal's] soul and body were designed and constructed by the supreme and invisible Demiurge, but the motions of the celestial system belong to the stars that constitute that system, and their astronomical movements manifest the activity of the cosmic soul itself. The cosmic animal is a *natural* being – one that has the status of a great god because its life is intelligent, immortal, completely self-sufficient and all-containing, but nonetheless natural because once it has been made by the Demiurge, being assigned its proper motion by him, it works on its own physical and psychological principles. With the mortal animals it is even clearer that they have been constructed to live their own lives in accordance with their own natures and that of the environment. The ancillary gods fashioned their complex respiratory and metabolic systems, but it is clearly the mortal animals *themselves* that will be breathing in and out and undergoing the stages in the metabolic cycle. [...] In short, the animals will be living their own lives. (Broadie 2011: 258)

Broadie's treatment of the issue is interesting because it neatly divides the roles of mind and necessity between that of creating the structure of the universe and its constituents, and that of the processes in a universe under this structural constraint. Once appropriately structured, Nature becomes autonomous from its divine origin, although it still bears the mark of intelligence in the aptness, the closure and the proportion of its structure. But Broadie also goes on to alert a serious problem for Plato's account in the *Timaeus*, namely that of *reproduction*. She indicates that Plato's account is attractive for the anatomical and physiological facts about existing living beings, but that it fails when the power of producing new structures through reproduction is concerned:

The activity [or reproduction] is said to be a *forming of the unformed*, and herein lies the philosophical difficulty. There is no place for this sort of mindless formation in the great dualistic scheme of Intelligence and Necessity. The domain of Necessity can only submit to formation of the organic structures, not initiate or guide it. And Intelligence goes to work through *intellectual* activity such as the divine demiurgic thinking that gives itself a goal reflectively, works out ways to meet it, takes note (a kind of reflection) of the beauty of the product, makes reasoned decisions prioritising one value such as intelligence in mortals over another such as longevity; or such as the thinking of the cosmic soul that continuously articulates relations of sameness and difference so as to produce in itself understanding (*nous*), knowledge, and firm and true beliefs and convictions; or such as the thinking of mortals, whose intelligence too is intended for truth about identities and differences, but which keeps staggering into confusion through interference by alien impulses, until rescue comes via the deliberate self-care of a

*learning*-regime focused on the ‘harmonies and revolutions of the All’. Moreover, all intelligence is immortal. The point is mentioning these features of Timean intelligence in different examples is to bring out the fact that in all these ways the forces that supposedly shape the embryo could hardly be more different. They are mortal psycho-physical forces crafted by the secondary demiurges. So far as animals are subject to these forces they are not reflective; they do not deliberate; their operation is not spelt out by reference to truth, knowledge, or learning; their fulfilment is defined not as self-perfecting assimilation to the cosmic harmonies and revolutions of the unique and everlasting cosmos, but as the successful procreation of mortal by mortal. Yet in carrying out their work these forces presumably do what so far only divine demiurgy guided by the extra-mundane paradigm has been shown as able to do, namely persuade material Necessity to contribute to their project. (Broadie 2011: 270-271)

The conclusion thus seems to be that, in order to account for the generation of animals, Plato would have to allow either that generation always requires a transcendent intervention, thereby losing the idea that nature can take its ordinary course once created, or that the structures of animals can arise by material necessity alone, thereby losing the explanatory role of intelligence in physiology and anatomy. This is a familiar dilemma, since it already came up in my Chapter 2 discussion of the problem of animal generation in Early Modern philosophy. Here, we find it prefigured in Plato. As we will see, this is not coincidental, since most Medieval and Early Modern thinkers would adopt the Platonic dualism of demiurge and nature, of Intelligence and Necessity. In doing so, however, they opted against the Aristotelian attempt to escape the dilemma by invoking *natural* teleology.

Aristotle followed Plato in the assumption that proper explanation in physics requires reference to teleology. This is clear, for instance, from his criticism of Anaxagoras in *Metaphysics A*, which clearly echoes the one offered by Socrates/Plato in the *Phaedo*. Like Socrates, Aristotle expresses initial appreciation for Anaxagoras and his principle:

When these men [i.e. the earlier presocratics] and the principles of this kind [i.e. the formal, material and efficient causes] had had their day, as the latter were found inadequate to generate the nature of things, men were again forced by the truth itself, as we said, to inquire into the next kind of cause. For surely it is not likely either that fire or earth or any such element should be the reason why things manifest goodness and beauty both in their being and in their coming to be, or that these thinkers should have supposed it was; nor again could it be right to ascribe so great a matter to spontaneity and luck. When one man said, then, that reason was present – as in animals, so throughout nature – as the cause of the world and of all its order, he seemed like a sober man in contrast with the random talk of his predecessors. We know that Anaxagoras certainly adopted these views, but Hermotimus of Clazomenae is credited with expressing them earlier. Those

who thought thus stated that there is a principle of things which is at the same time the cause of beauty, and that sort of cause from which things acquire movement. (984b)

Within pages of this praise, however, Aristotle repeats the Socratic judgement by stating that “Anaxagoras uses reason as a *deus ex machine* for the making of the world, and when he is at a loss to tell for what cause something necessarily is, then he drags reason in, but in all other cases ascribes events to anything rather than reason” (985a). However, he also goes on to suggest that Plato’s response is not satisfactory either, since it is confined to the material and the formal causes (988a). I find it plausible that we should take Aristotle’s closing judgment on the issue as equally targeted at Plato:

That for the sake of which actions and changes and movements take place, they assert to be a cause in a way, but not in this way, i.e. not in the way in which it is its *nature* to be a cause. For those who speak of reason or friendship class these causes as goods; they do not speak, however, as if anything that exists either existed or came into being for the sake of these, but as if movements started from these. In the same way those who say the One or the existent is the good, say that it is the cause of substance, but not that substance either is or comes to be for the sake of this. Therefore it turns out that in a sense they both say and do not say the good is a cause; for they do not call it a cause *qua* good but only incidentally. (988b)

This passage is surely a difficult one, but I find it highly plausible that in it, Aristotle is claiming that even those among his predecessors who aimed to take final causes into account failed to do so because they fashioned it after the model of another cause. For Aristotle, it is specific to a final cause that we explain a process by referring to its end state. Most thinkers, however, seem to place the good or the rational in the initial state. This was, of course, also what Plato did in the *Timaeus*: the teleology exhibited by the universe was due to its being fashioned after the intelligible paradigm intended by the divine artisan, the demiurg. By thus equating the intelligent cause with the paradigmatic pre-existence, Plato is effectively reducing, in Aristotle’s opinion at least, final cause to formal cause.

D. M. Balme expresses the import of Aristotle’s departure from Plato on the issue of teleology excellently in the following passage:

The novelty in Aristotle's theory was his insistence that finality is within nature: it is part of the natural process, not imposed upon it by an independent agent like Plato's world soul or Demiourgos. This is what allows him to claim that none of his predecessors had recognized the final cause with any clarity. Anaxagoras called his primary cosmological cause 'Mind', and for this Aristotle likened him to a lone sober man among drunks; Plato offered cosmic teleological causes in the *Timaeus*, *Philebus* and *Laws*; Xenophon argued for the popular belief in providential

guidance of natural phenomena. But such constructions are not what Aristotle meant by the final cause. Nor has his natural teleology anything to do with intentionality, the physiology of which in man and animals he explains in *MA*. There is no deliberating or purposing in most animals, he says; and it is by nature alone that roots and leaves grow for the sake of fruit. (Balme 1987: 275)

The major difference that Aristotle saw between his own approach and that of others is that he was focused on understanding nature and natural change as opposed to artificial change. This is born out well by his famous analysis of the notion of nature in Book II of the *Physics*. His famous definition of nature is there that of the “internal principle of motion and rest” (192b). By this, he means that we can and do recognize natural things and natural motions because they somehow flow from the fact that this thing is this particular kind of thing, and no other. The upshot of this is that nature, for Aristotle, is both autonomous or self-moving and strongly linked to the idea of being a specific kind of thing.

When Aristotle stresses the distinction between natural entities and artifacts, he insists that they differ in that the former are themselves causes of the changes they normally and therefore regularly<sup>4</sup> undergo, whereas artifacts do not have such an internal principle of change and are entirely dependent on the aid of the artificer. The important factor here is that, for Aristotle, change is brought about by and within nature, and not (normally or regularly) externally imposed upon it.

It becomes clear from the concept of nature as it is characterized in *Physics* II.1 that Aristotle is thinking of self-change primarily in terms of self-production, more specifically as the self-production exhibited by organisms in their process of genesis:

Nature [φύσις] in the sense of a coming-to-be [γένεσις] proceeds towards nature. For it is not like doctoring, which leads not to the art of doctoring but to health. Doctoring must start from the art, not lead to it. But it is not in this way that nature is related to nature. What grows *qua* growing grows from something into something. Into what then does it grow? Not into that from which it arose but into that to which it tends. The shape then is nature. (193b)

Aristotle clarifies here what he means when he says that form is nature, a thesis he advanced against the materialist thesis that the underlying matter of things is their nature, and that natural change should be understood primarily in terms of the changes necessitated by the basic elements and material principles. This clarification is necessary because there is a familiar sense in which things can be determined by their

---

<sup>4</sup> Aristotle believes that we can rather easily identify the proper motions of natural kinds by seeing how they regularly act in non-contrived situations. The challenge of the natural philosophers consists in identifying the natures and the principles that underlie these regular motions. (Cf. Waterlow 1982: 33-34)

form without this form thereby being something distinctly causally active over and above the matter of which it is the form. This familiar sense is that of the artifact, where the artificer draws upon the natural properties and behaviours of his materials to bring about certain effects. The form itself, however, is not distinctly causally active in such a case. If Aristotle wants to defend his idea that the form is distinctly operational apart from the matter, he has to reveal a strong disanalogy between the natural and the artificial. He tries to do this by insisting that, in natural beings, the form is causally active because it leads the process of production, of coming-be. Natural beings are self-producing because they have an internal tendency to realize in themselves a form that they do not yet fully have. This refers of course to the doctrine of final causation: for Aristotle, forms are distinctly causally relevant because they serve as final causes.

To appreciate what this means, we need to look at the case of doctoring mentioned by Aristotle. In the case of doctoring, the end of the activity and its principle of change are not the same: a doctor's activity is best described<sup>5</sup> as exercising the art of healing, not as becoming-healthy. The generation and growth of a given plant, however, is best described as becoming that plant, not as an exercise of plant-making. The difference between the two cases, for Aristotle, is that in the first case something other than the form is bringing about the form, whereas in the latter case the form is bringing about itself.

Another important aspect of Aristotle's theory is that the forms of natural things refer to the *kind* of thing. Specifically, they are substantial forms in that they characterize that of which they are the form as one substance. Material principles, Aristotle believed, cannot make sense of the fact that, in nature, there are things best understood as existing and persisting individuals of specific kinds. We would now tend to say that this is because materialism explains natural things in mass-terms, which do not allow for individuation, and not in terms of sortals, which serve to identify individuals.<sup>6</sup> For Aristotle, this is problematic mostly because he believes we need to admit the individuality and the kind of a thing to be causally active in the world.

This is not to say that Aristotle leaves no room for necessity and material principles. On the contrary, he believes that we need to understand how, in the processes of genesis that give rise to natural individuals, the material necessity is employed as a means and a constraint. As John Cooper expresses it:

---

<sup>5</sup> It is a central tenet of Aristotle's *Physics* that there are better and worse descriptions for physics, not just for convenience sake, but principally. A situation exhibits its causal goings-on only under a relevant description, not under an irrelevant, "accidental" one.

<sup>6</sup> See also 5.4.2.

according to Aristotle in *Physics* II.9, Democritean necessity does play a role in nature, i.e. in the formation of natural, living things, but it does so only hypothetically - that is, only on the hypothesis that a living thing is to be produced. *Given* that a living thing is to be, certain materials are necessary (i.e. these materials are hypothetically necessary); this means that the presence of those materials is to be explained by reference to this goal. But once, for that reason, they *are* there, their nature, and the material necessity that belongs to it, will cause them to behave in various ways. They will act in those ways by necessity, and this necessity is not a hypothetical one.

[...] [Aristotle] wants to say that Democritean necessity does indeed exist and has a role to play in the formation of living things, but that where it does make a contribution it only does so because the materials whose necessary action is in question are themselves necessitated hypothetically. On this interpretation it could be said that Aristotle 'subsumes' Democritean necessity under hypothetical necessity, in the explanation of living things. But that does not mean that he *reduces* it to hypothetical necessity. (Cooper 1987: 263-264)

This position should remind us of Plato's solution, in the *Timaeus*, to the allocation of roles to mind and necessity. There, mind was considered a primary cause, whereas necessity constituted necessary causes. Like Plato, Aristotle also stresses that the more important focus of natural philosophy should be on form, not on matter and necessity.

But these similarities threaten to cover up another important feature of Aristotle's theory. For Plato, forms determined natural things because they were imposed upon them by an artificer. On Aristotle's definition, however, this would make natural beings into artifacts, whose proper cause is the art of the artificer, not their own natures. Most importantly, it would be impossible for natural things to come about naturally. As we discussed above, Plato's demiurgic approach ran into difficulties in the explanation of reproduction, of the tendency of natural things to bring forth natural things of their kind naturally and unintentionally. It is revealing that Aristotle seemed to define nature precisely as genesis, as this self-production and coming to be with which Plato struggled.

This brings us to the most puzzling and fascinating aspect of Aristotle's philosophy of nature, namely the idea that a final cause in nature is not usually intentional. The final cause does not give rise to its end because it is prefigured as a conscious goal of the agent. For Aristotle, this is born out perfectly by the case of animals, which often act in complex concerted manners without having the rational capacities to plan these actions, and more importantly by plants, which reveal, in their growth, nutrition and generation, a finality at work without even the semblance of intentionality.

The doctrine of natural teleology, which holds that the tendency to realise ends is a property of nature itself, and not one superimposed by a transcendent function, is defining for Aristotle's biology in one more important sense, namely in the idea of a vegetative soul. The vegetative soul is not a soul in any sense familiar to us, because it



confers upon its substance neither intention and reason, neither sensation nor appetite, but only the capacity for growth, nutrition and generation, i.e. the capacity for self-production.

In subsequent centuries, however, it was Plato's account, not Aristotle's, which was to define our understanding of final causation. For philosophers from Jewish, Christian or Islamic backgrounds, the *Timaeus* offered a picture of the relation between intelligence and necessity that was far more conducive to their religious views, one that showed us a world which bore the mark of intelligent design in the purposiveness of its structure and its constituents. Aristotle, however, gave us an autonomous world filled with things that are responsible for their own natures, depending on nothing else for their purposive alignment. This innovative move by the Stagirite smacked of the glorification of nature that later authors would identify as idolatry; as Dennis Des Chene remarks (commenting on the passage from David Balme I also quoted above: "[t]hat innovation is rescinded in Christian Aristotelianism, no doubt because the Christian God provides a counterpart to the Demiurge, and because it was fitting that the nonhuman world should depend on God as an instrument on its maker" (Des Chene 1996: 187).

More telling than the obvious Platonic leanings of Philo or Augustine and those they inspired is, however, the fact that the infamous Medieval Aristotelians never adopted the doctrine. For many of them, the idea that end states can influence earlier states is ludicrous, except in the case of intentionality:

It is thus generally valid for scholasticism – and it is characteristic that this principle is never subjected to doubt – that everything that happens in the world is caused by a *causa efficiens* (*proxima* or *remota*), that acts for an end, and conversely: that every *finis*, insofar as it works as a "cause" (and is not simply the end [Ziel], the *terminus*, of an *actio*) presupposes an *agens*, that immediately or mediately consciously strives after the endpoint [Ziel] as a goal [Zweck]. A direct, as it were *a fronte* working, "attracting" (and not "driving") causality, through which the end determines its means and in a way urges the *agens* to work by means of a kind of reversed efficient causality, is never taken into consideration by scholasticism. A *finis* is always an end strived after by some will, that can only work as a cause *ut apprehensum*. (Maier 1955: 278; my translation)

The scholastics therefore sought to save Aristotle from this error by developing the idea that the finality exhibited by non-intentional natural processes is not to be sought in the anticipation of the form by the natural thing, but in the intention God had for the activities of that thing (Cf. Maier 1955: 277). On this account, nature itself is not teleological – it merely exhibits the intentionality of God. As Aquinas expressed it in his commentary on the physics: "From this it is clear that nature is nothing else than some form of art, i.e. a divine art, impressed on things, by which these things are moved to a determinate end (Aquinas 1954: 268; my translation). In the hands of theologians of

various persuasions, nature had become an artifact, impotent to achieve its own purposes or even have them.

#### 4.1.2 The Mechanization of the World-Picture as the Artifactualization of the World

In the standard historiographies of philosophy, we are usually regaled with the tale of the heroism of early modern thinkers, who replaced a thoroughly unscientific world-view, that of the scholastics, which was concerned not with analyzing real causes, but with investigating the theological stamp on nature, with a scientific world-view, in which men awoke to find the universe bereft of providence and ruled by blind necessity. But as usual, the standard historiographies of an epoch say more about the describing era than about the described era. If the likes of Descartes were capable of refocusing the study of nature on the efficient and material causes rather than on the final causes and the (kinds of) formal causes implied by them, this was because scholasticism had already thoroughly prepared this shift. As Anneliese Maier has shown, the idea that final causes were only causes in a derivative sense, and therefore inadmissible in natural philosophy, was already well established and elaborated in the fourteenth century. It is on this achievement that Descartes based his notorious and seminal banishment of final causes from natural philosophy:

When dealing with natural causes we will, then, never derive any explanations from the purposes which God or nature may have had in view when creating them <and we shall entirely banish from our philosophy the search for final causes>. For we should not be so arrogant as to suppose that we can share in God's plans. We should, instead, consider him as the efficient cause of all things. (AT VIII.1: 15-16; Descartes 1985: 202)

From one perspective, this passage is not nearly as revolutionary as is usually assumed. The idea that the consideration of final causes is improper in natural philosophy was shared by many scholastics, even though they of course still chose to understand causality in terms of substantial forms. For Descartes, if formal cause could enter into considerations of natural philosophy at all, it would take a very different form indeed.

Don Garrett has suggested that Descartes's rejection of final causation is based on the fact that "[w]ill, operating through acts of volition, is the only teleological selection process that he acknowledges" (Garrett 1999: 326). He goes on to connect this with Cartesian dualism:

In fact, Descartes's ontological dualism of extended and thinking substances seems to be reflected in a dualism of types of explanation. Extended substances, he implies, cannot themselves produce effects directly through teleological

selection, whereas thinking substances cannot directly produce effects mechanistically. Because thinking substances and extended substances can interact causally, in Descartes's view, a given state of matter may be explained teleologically (as when a muscle moves in order to move an arm), whereas a given state of mind (such as a physiologically induced passion) may be explained mechanistically. Of course, the alleged unintelligibility of Cartesian mind-body causal interaction may call into question the ultimate adequacy of these explanations. But in any case, whereas each kind of Cartesian substance may participate in either or both kinds of *explanandum*, each can evidently participate only in its own proper kind of *explanans*, each grounding or originating only its own proper kind of explanation. Minds, but not bodies, can have ends, goods, or purposes; and bodies, but not minds, can have physical structures and distributions of forces. (Garett 1999: 326)

If this is right, we can see Descartes reverting to the position suggested by Plato's *Timaeus*, namely a dualism of mind and necessity, where mind is solely responsible for the teleological, and necessity governs all interactions. Since Descartes is convinced that we need no recourse to teleology to explain any natural occurrences that do not involve human intentional and conscious activity, he can completely omit consideration of teleology in natural philosophy. But this does not in itself mean that teleology is out of the picture. Like Plato and the Christian Aristotelians, Descartes could have supposed that the operations of appropriately structured matter need not be explained through final causation, but that their structure, and the appropriateness of this structure, needs to be explained by referring to divine intentions.

It is on this point that Descartes departed from the tradition. For him, it was impossible to properly investigate the possible uses of the structure of the material world. Since Descartes's God has an intelligence that is so vastly removed from the intelligence we know from our own case, we can make no reasonable analogies concerning his intentions. Nor need we: we can perfectly take the structures in the world as given, as brute facts, and then examine these structures and the way in which they constrain mechanical interactions. The peculiarity in Descartes's world-view is, then, that we continue to focus on natural entities from the perspective of theology, namely insofar as they are artifacts and instruments of a divine artificer, but omit reference to this artificer altogether. This sleight of hand, I fear, has fooled many into believing that Descartes's mechanical universe is any less a theological world-view than that of his scholastic predecessors.

It did not fool the mechanical philosophers themselves, however. As mechanism left Descartes's hands and entered what we now like to regard as the scientific revolution, the banishment of final causes did not become a fundamental tenet of the

mechanical philosophy. In fact, they were ultimately reintroduced, with their full theological import, by Robert Boyle.<sup>7</sup>

Boyle responded to Descartes's dismissal of reference to final causes in natural philosophy in a variety of ways. He argues for instance that it is not presumptuous, but rather prudent, to infer that some of the most obviously contrived entities in the world, such as the eye, are created with their most obvious use at least partly in mind. Besides this, he asserts that the consideration of final causes in such a case does not in any way impede the properly physical way of understanding the entity:

I see not how God's designing some of his works for particular uses, amongst others, is inconsistent with the physical accounts of their creation. Thus, a man may give a mechanical reason of the structure of every wheel, and other part of a watch, and of their way of acting upon one another, when rightly set together, and, in short, of the contrivance and phenomena of the little machine; tho' he supposes that the artificer design'd it to shew the hours of the day, and tho' he has that use in view, whilst he explains the fabric and operations of the watch. (Boyle 1725: 155-156)

In this passage, Boyle is making a distinction between structure and function, and argues that there is no difficulty in considering both, since they do not infringe upon each other. The function of a mechanical system can be cited in order to argue *why* it was contrived, and its structure may suggest *that* it was contrived, but attribution of function does not run counter to description of structure.

The account just presented seems to suggest a neat division of labour between physics, in which we consider the structure and mechanical causes of physical things, and natural theology, in which we are concerned with the final causes of things. The reason why these two kinds of accounts can be neatly distinguished is the same as in the case of some 14<sup>th</sup> century scholastics, and of Descartes, namely the rejection of unintentional, i.e. natural teleology:

---

<sup>7</sup> I suspect that Boyle was inspired here, as in other things, by Francis Bacon, who wrote in *The Advancement of Learning*: "in the work of the creation we see a double emanation of virtue from God; the one referring more properly to power, the other to wisdom; the one expressed in the subsistence of the matter, and the other in disposing the beauty of the form." (Bacon 1876: 295) Bacon suggests here that the two aspects are not only different aspects of creation, they also correspond with different divine attributes. In the subsistence of matter, we find the God's power to create and maintain existents, and in its structure or form, we find his intentional disposing of matter. In later naturalist accounts, the power of God would be transferred, giving nature itself the power to subsist and matter itself its general properties. The correlate of wisdom, form, however, would never be conveniently naturalized, and would remain, to this day, an obstacle to any naturalist project based on mechanism.

There are two accounts on which the actions of natural agents may be said to tend to a certain end; as either when the agent has a knowledge of that end, and acts with an intention to obtain it; or, when the action of the proximate agent is directed as it ought, to obtain an end, which, yet, is neither known, nor intended, by that proximate agent, but by a remoter, that is intelligent. In the former of these senses, I cannot admit, that any inanimate body acts for ends, since that presupposes the agent both to know the end he is to obtain, and to purpose to obtain it: things whereof inanimate bodies are incapable. And, to fancy with some, that they may have a knowledge, *sui generis*, as they speak; which, tho' confined to the actions proper to a particular kind of body, shall yet suffice to determine them to those actions, is, to offend against that rational, and receiv'd rule in philosophy, that Beings are not to be multiply'd, without there appears a necessity for them; and to introduce a sort of knowledge that seems unintelligible.

It remains, then, that I embrace the second sense, in which 'twas formerly said, natural things may work for an end, tho', in this case too, we must speak somewhat improperly, for the action may more justly be attributed to the remote intelligent, than to the immediate agent, which is but, as it were, the instrument of the other. Now, it appears to me, that the most wise, and powerful author of nature, whose piercing sight is able to penetrate the whole universe, and survey all the parts of it at once; did, originally, frame material things into such a system, and settle among them such laws of motion, as he judged suitable to the ends that he proposed to himself in making the world. (Boyle 1725: 170)

This passage will be important for our understanding of Kant's concept of purposiveness. First, Boyle rejects natural teleology because it violates Ockham's razor. Indeed, as we saw, the doctrine of natural teleology was intimately connected, in Aristotle, with the idea that we need to postulate distinct natural kinds of things in order to understand nature. Aristotle's insistence on the explanatory relevance of species seems a clear violation of the economical considerations of natural philosophy. For Kant, this connection would prove essential, as we will see in the next chapter, since he recognized that the mechanical philosophy and the idea of ontological and explanatory parsimony are strongly connected. Second, Boyle suggests that final causation in nature may be incomprehensible when we limit ourselves to finite intellects, but that it becomes unproblematic when we ascribe the intentions in nature to an infinite intellect transcendent to nature and capable of regarding it entirely at once. We will address this picture again later in this chapter, when we discuss the concept of the supersensible in Kant's "Antinomy of Teleological Judgment".

The picture offered just now seems to suggest a neat division of labor between natural philosophy, the proper business of which is the explanation of natural phenomena on the basis of mechanical descriptions of structures, and natural theology, which is concerned more with the origins of and reasons for these mechanical structures. But the division is not as neat as would appear at first sight, because Boyle

admits that there is a way in which considerations of final causation, of function, can be of use in physics proper:

There are two ways of reasoning from the final causes of natural things, that ought not to be confounded. For sometimes men draw arguments from the use of bodies, that relate to the author of nature, and the general ends he is supposed to have intended in things corporeal, as when from the manifest usefulness of the eye, and all its parts, for vision, 'tis inferred that the eye was originally framed by a very intelligent Being, with a particular care that animals should be furnish'd with the fittest organ of so necessary a sense. And sometimes, also, men ground arguments upon the supposed ends of things, as to the peculiar nature of things themselves, and conclude, that this affection of a natural body or part ought to be granted, or that deny'd, because by this, and not by that, or by this more than by that, the end design'd by nature may be best and most conveniently obtain'd. The latter sort of arguments I usually call purely physical, and those of the former may be called physico-theological, or, by a shorter name, metaphysical. (Boyle 1725: 173)

Boyle thus allows for reference to final causes in natural philosophy as well because they can guide our investigation into mechanical structures. Even in this case, however, he tends to restrict their use to a very distinct kind of investigation. He advises against using considerations of final causation in astronomy, and admonishes us to employ them only when reflecting on the nature of animals (by which he also means plants). Moreover, even in the case of animals, Boyle will insist that we never take into account their uses to other animals, but only consider “those ends and uses of the parts of an animal, that relate to the welfare and propagation of the animal itself, and which, therefore, I call animal ends” (Boyle 1725: 177). Finally, he claims only that such considerations will help us in our investigations into structure by allowing us to “draw probable conjectures” (Boyle 1725: 177).<sup>8</sup>

Boyle's account of teleology partly became influential because it cleared the ground for the two uses of final causation just described. On the one hand, they allow us to infer, from the mechanical structures of the universe, the existence of a supreme mechanic responsible for these structures. In this way, he spawned the British tradition of natural theology, which sought to forge or strengthen the alliance between the mechanical philosophy and theology. It is not coincidental that Boyle's testament arranged for lectures to be held on these topics, the so-called Boyle lectures. On the

---

<sup>8</sup> It is in this context that Boyle makes his famous ascription to Harvey of the claim that it was considerations on the uses of the valves in the veins that allowed him to come up with the idea of the circulation of the blood. (Boyle 1725: 179-180).

other hand, Boyle allowed for a heuristic use of final causation in natural philosophy. On this interpretation, final causes had no explanandum of their own, no explanatory role to play in natural philosophy, but could be adopted in order to facilitate the search for proper mechanical descriptions and explanations. It is against these two legacies that Kant would react in the “Critique of Teleological Power of Judgment”.

## 4.2 Natural Purpose and Circular Causation

### 4.2.1 Natural Purposiveness Delineated

Kant’s most important discussion on topics that we would now identify as belonging to the philosophy of biology, is to be found in the second part of his 1790 *Critique of the Power of Judgment*, which he named the “Critique of the Teleological Power of Judgment”. This part is again composed of three major divisions, or rather two divisions and a lengthy appendix. The first division, the “Analytic of the Teleological Power of Judgment”, is devoted to the analysis of the concept of a natural purpose and the concomitant idea of teleology or purposiveness that is one of *objective, material* and *intrinsic* purposiveness. In this subsection, I will discuss the way in which Kant contrasts this form of purposiveness with various others. In the next two subsections, I will discuss his example of a natural purpose and his elaboration of the concept of a natural purpose respectively.

The first important contrast Kant introduces is that between subjective and objective purposiveness. Subjective purposiveness was the topic of the first part of the *Critique of the Power of Judgment*, the “Critique of the Aesthetic Power of Judgment”. There, Kant argued that

[o]ne has good reason to assume, in accordance with transcendental principles, a subjective purposiveness of nature in its particular laws, for comprehensibility for the human power of judgment and the possibility of the connection of the particular experiences in one system of nature; where among its many products those can also be expected to be possible which, just as if they had actually been designed for our power of judgment, contain a form so specifically suited for it that by means of their variety and unity they serve as it were to strengthen and entertain the mental powers (which are in play in the use of these faculties), and to which one has therefore ascribed the name of **beautiful** forms. (AA V: 359)

Subjective purposiveness is therefore the purposiveness of the structure of nature for our cognitive capacities and powers. In the “Critique of the Aesthetic Power of

Judgment”, Kant had already advanced his idea that an objective purposiveness requires that there be a determinate end, and thus a concept, in light of which we judge the purposiveness of the object, whereas in the case of subjective purposiveness, no determinate concept is involved in the aesthetic judgment (AA V: 227-228). This contrast serves us fairly little here, because the notion of subjective purposiveness is far less readily understood, far more specific to Kant’s own system, than that of objective purposiveness. We will therefore only return to the notion of subjective purposiveness in the last chapter of this dissertation. For now, it suffices to note that objective purposiveness contrasts with subjective purposiveness because the former involves a judgment on the object, and more precisely the nature of that object itself.

The objective purposiveness of which Kant intends to treat in the “Critique of Teleological Power of Judgment” is, moreover, material rather than formal. Formal objective purposiveness is the purposiveness exhibited by mathematical principles, figures, etc.:

All geometrical figures that are drawn in accordance with a principle display a manifold and often admired objective purposiveness, namely that of serviceability for the solution of many problems in accordance with a single principle, and indeed of each of them in infinitely many different ways. The purposiveness here is evidently objective and intellectual, not, however, merely subjective and aesthetic. For it expresses the suitability of the figure for the generation of many shapes aimed at purposes, and is cognized through reason. But the purposiveness still does not make the concept of the object itself possible, i.e., it is not regarded as possible merely with respect to this use. (AA V: 362)

The idea of formal purposiveness, and its placement in the discussion, is somewhat peculiar to contemporary readers. In many ways, it is a digression that is no longer central to our thinking on the subject, but it is important to discuss it here precisely for the misunderstandings it may otherwise engender. The inclusion of the idea of formal objective purposiveness here is the direct result of Kant’s criticism of traditional, Pythagorean-inspired aesthetical theories that believe certain geometrical figures (or numbers, for that matter), to be aesthetically pleasing in virtue of their mathematical properties. If such theories were right, then Kant’s own aesthetics would be flawed, for these figures would be beautiful in virtue of the fact that they are “mere representations of a determinate concept” (AA V: 241), and Kant’s theory holds that aesthetic judgment cannot include reference to *any* determinate concept. He therefore argues that such approaches are mistaken, and advances instead that we appreciate these geometrical shapes for their usefulness in mathematics:

It is customary to call the properties of geometrical shapes as well as of numbers that have been mentioned **beauty**, on account of a certain *a priori* purposiveness, not expected from the simplicity of their construction, for all sorts of cognitive



use, and to speak of this or that **beautiful** property of, e.g., a circle, which is discovered in this way or that. But it is not an aesthetic judging by means of which we find it purposive, not a judging without a concept, which makes noticeable a merely **subjective** purposiveness in the free play of our cognitive faculties, but an intellectual judging in accordance with concepts, which gives us distinct cognition of an objective purposiveness, i.e., serviceability for all sorts of (infinitely manifold) purposes. One would have to call it a **relative perfection** rather than a beauty of the mathematical figure. (AA V: 366)

For Kant, then, the purposiveness exhibited by geometrical shapes and by numbers is objective rather than subjective. But this purposiveness is not thereby to be attributed to the concrete objects that instantiate these shapes or numbers, for the familiar reason that, in mathematics, it is these shapes and numbers which concern us, and (somehow) not the concrete empirical objects which happen to be characterized in mathematical terms. For this reason, Kant indicates that formal purposiveness is *relative to* a certain cognitive enterprise of man, namely the employment of mathematics for a variety of uses.

This brings us to the final major contrast that Kant lists, namely that between relative and internal purposiveness. A useful initial characterization is that an object is relatively purposive insofar as it is a means for the end of another object, and internally purposive insofar as it is a means to its own ends. This initial characterization will require much more comment – in a way, Kant’s entire analysis of natural purposiveness is intended to bear out what it means to be internally purposive.

In discussing relative purposiveness, Kant typically refers to either ecology or human activity. In such a way, oxen are purposive as food for carnivores, or purposive as beast of burden for humans. But Kant protests that it is extremely problematic to ascribe such purposes to objects, because they do not characterize the objects themselves, and are rather external to them. There is thus no specific need, in understanding these objects, to take into account their usefulness to other objects.

But there is an initial problem in this picture, one that becomes clear once we reflect on an important similarity between an ecology and an organism. Kant seems to assume that relative purposiveness is the subordination of one thing to another as a means. This is born out by his central argument against the consideration of relative purposiveness:

From this it can readily be seen that external purposiveness (advantageousness of one thing for another) can be regarded as an external natural end only under the condition that the existence of that for which it is advantageous, whether in a proximate or a distant way, is in itself an end of nature. This, however, can never be made out by mere contemplation of nature; thus it follows that relative purposiveness, although it gives hypothetical indications of natural ends, nevertheless justifies no absolute teleological judgments. (AA V: 368-369)

In this passage, Kant assumes that external purposiveness means a subordination of an object to the ends of another, for only on that assumption need there be a member in the chain of sub- and superordinated means and ends that is an end in itself in order for the relatively purposive to be legitimately regarded as purposive at all. But this assumption is not necessary for an assessment of an ecology, for we might regard ecologies as self-sustaining due to the reciprocal means-ends relations obtaining between their various subsystems. Such a “relative purposiveness” could not be dismissed on the given basis, since it is precisely the purposiveness that obtains between the various relevant parts (organs in the most general sense of the word) of an organism. Initially, then, it would seem that we need either accept that ecologies exhibit teleology, contrary to what Kant seems to say, or deny that we can legitimately ascribe purposiveness to organs, which Kant explicitly wants to do.

I will leave the issue of relative purposiveness hanging here until the conclusion of the next subsection. Only a better understanding of Kant’s concept of internal purposiveness and natural purposiveness will allow us to resolve this puzzle. I have raised it, however, because I do not believe clarity would be gained by taking Kant’s concept of relative purposiveness at face value. On the contrary, this misgiving will be useful in our further analyses, first of all in our assessment of Kant’s most important illustration for the concept of a natural purpose.

We have seen now that, in §§61-63, Kant sought to negatively delineate the notion of purposiveness with which he will be concerned in the “Critique of Teleological Power of Judgment”, from subjective purposiveness, from objective formal purposiveness, and from objective material relative purposiveness. The last distinction is the most important one for the present discussion, since it implies that objective material internal purposiveness characterizes a purpose not as relative to some external end, but rather as relative to an internal end. This raises the question, of course, of what it means for a natural object to have internal ends, or, as Kant says it, to *be* an end:

In order to see that a thing is possible only as an end, i.e., that the causality of its origin must be sought not in the mechanism of nature, but in a cause whose productive capacity is determined by concepts, it is necessary that its form not be possible in accordance with mere natural laws, i.e., ones that can be cognized by us through the understanding, applied to objects of the senses, alone; rather even empirical cognition of their cause and effect presupposes concepts of reason. Since reason must be able to cognize the necessity in every form of a natural product if it would understand the conditions connected with its generation, the **contingency** of their form with respect to all empirical laws of nature in relation to reason is itself a ground for regarding their causality as if it were possible only through reason; but this is then the capacity for acting in accordance with ends (a will); and the object which is represented as possible only on this basis is represented as possible only as an end. (AA V: 369-370)

In the first sentence of this passage, Kant clarifies that a thing that is also an end is not possible through mere natural laws because its concept determines its production. Moreover, he specifies that we require not just concepts of the understanding, but also concepts of reason, to properly cognize such an object. The claim is thus that, on the one hand, mechanism, understood as explanation according to natural laws alone, does not suffice for explaining the object understood as end, because it underdetermines the form of the object (hence the reference to the contingency of this form with respect to natural laws), and that, on the other hand, such an object prompts us to adopt the idea of causation determined by a concept of reason, i.e. the causation of will. In order to make this idea clearer, he gives an initial example:

If someone were to perceive a geometrical figure, for instance a regular hexagon, drawn in the sand in an apparently uninhabited land, his reflection, working with a concept of it, would become aware of the unity of the principle of its generation by means of reason, even if only obscurely, and thus, in accordance with this, would not be able to judge as a ground of the possibility of such a shape the sand, the nearby sea, the wind, the footprints of any known animals, or any other nonrational cause, because the contingency of coinciding with such a concept, which is possible only in reason, would seem to him so infinitely great that it would be just as good as if there were no natural law of nature, consequently no cause in nature acting merely mechanically, and as if the concept of such an object could be regarded as a concept that can be given only by reason and only by reason compared with the object, thus as if only reason can contain the causality for such an effect, consequently that this object must be thoroughly regarded as an end, but not a natural end, i.e., as a product of **art**. (AA V: 370)

The example here is that of one stumbling across a perfect simple geometrical figure in nature. In such a case, Kant has it, one would be immediately tempted to assume this figure to have been made by human hands. Unfortunately, Kant's own presentation of the reasoning that would lead one to such a conclusion is somewhat enthymematic, since he suppresses important premises. The most important premise is that, in nature, we usually do not find regular geometric shapes: such shapes are considered and constructed by men as idealizations of the complex natural shapes in order to understand the latter. It would thus seem an extreme coincidence if natural laws would conspire to form such a regular shape. Given this extreme coincidence, Kant believes it is natural to assume that somehow, we must understand the concept of the regular geometric shape to have played a role in the generation of the figure. In labeling the figure as a product of art, we claim that the concept has in fact played a role in the generation of the figure, namely insofar as the concept was *intended* by the *maker* of the figure.

Immediately after giving this example, however, Kant qualifies its scope and importance by noting that his concern here is natural ends, not artificial ends:

But in order to judge something that one cognizes as a product of nature as being at the same time an end, hence a **natural end**, something more is required if there is not simply to be a contradiction here. I would say provisionally that a thing exists as a natural end **if it is cause and effect of itself** (although in a twofold sense); for in this there lies a causality the likes of which cannot be connected with the mere concept of a nature without ascribing an end to it, but which in that case also can be conceived without contradiction but cannot be comprehended. (AA V: 370-371)

Kant notes that the account given of the regular geometric shape in his preceding example cannot be simply transposed to what he calls natural ends on pains of contradiction. He does not specify which contradiction this would invite, but the subsequent discussion reveals that it is, in all likelihood, that of ascribing a non-natural cause (a volitional, intentional act) to a natural entity. This is in fact a contradiction, because one would seek to explain what it means to be a natural purpose by making its purposiveness non-natural. Kant solves this problem by characterizing a natural end as that which is cause and effect of itself. Such an account, he states, can avoid the contradiction between the object's status as end and the status as natural, but avoids it on the pain of making this concept *incomprehensible*. In the final subsection of this section, we will see exactly what Kant means by a solution that resolves the contradiction but introduces incomprehensibility. First, however, we will need to turn our attention to Kant's own example of a natural end.

#### 4.2.2 Natural Purposiveness Exemplified

In this subsection, I will discuss Kant's example of a natural purpose as he introduces it in §64 of the "Analytic of the Teleological Power of Judgment", which he intends as a preliminary characterization of the concept. This example, I believe, elucidates much more than is usually recognized, if only because it reveals rather than explicitates much of Kant's thinking on the subject.

In §64, Kant enumerates three ways in which a tree can be said to produce itself. These three ways are best characterized as reproduction, growth and/or nutrition, and regeneration. Before discussing the three separately listed characteristics as Kant presents them, I would first like to point out that they together seem to constitute the natural faculties, which were considered by Aristotle to be the powers of the vegetative Soul. For Aristotle, these powers were the minimal conditions for something to count as living (413b). It is important, moreover, to note that, for Aristotle, the presence of vegetative powers did not imply the presence of intentionality, since the latter is a property accompanying the *animal* soul, which needn't itself accompany the vegetative soul. This is also the reason why Galen (1916: 3) decided to dismiss talk of a vegetative

soul, since for him, soul implied (a modicum of) intentionality or consciousness, whereas the teleological properties of the natural faculties involved no intentionality. In choosing to talk of the vegetative properties, and of plants, Kant is thus referring to those properties which are traditionally regarded as non-intentionally teleological.

As Philippe Huneman (2007: 215-246) has noted, however, it is not obviously so that Kant regards plants as paradigmatically alive, for all his explicit definitions of the concept of life seem to refer to volition and intentionality, and exclude vegetative processes. This is already the case in a relatively early discussion of the topic to be found in the *Dreams of a Spirit-Seer Elucidated by Dreams of Metaphysics* from 1766, where Kant notes that “[t]he undisputed characteristic of life, belonging to that which we perceive by means of our outer senses is, doubtless, free movement, which shows us that it has originated from the power of the will” (AA II: 330). This passage seems to say that the concept of life is intricately linked to the concept of will, i.e. of free action. Moreover, in the context of this work, it also seems as if Kant is defining organic processes as material processes which seem to involve or exhibit the action of an immaterial principle, i.e. a spirit or soul. All of this would suggest that Kant is in fact identifying life and intentionality. Such a conclusion would be mistaken, however, since Kant immediately continues:

However, the conclusion that, when this characteristic mark is not encountered, then every degree of life is also lacking, is not certain. Boerhaave says somewhere: *The animal is a plant which has its root in its stomach* (inside itself). Someone else might, with equal propriety, play with these concepts and say: *the plant is an animal which has its stomach in its root* (outside itself). It is, therefore, possible for plants to lack the organs of voluntary motion and, in lacking them, to lack the external characteristic marks of life, which are certainly necessary to animals, for a being which has the instruments of its nourishment within itself must be able to move itself according to need. A being, on the other hand, which has the instruments of its nourishment outside itself and sunk in the element which supports it, is already adequately provided for by forces external to itself. Even though such a being contains within itself a principle of inner life, namely, vegetation, it does not need an organic arrangement to be made for external voluntary activity. None of this is necessary for my argument, for, apart from the fact that I should have very little to say in favour of such conjectures, these conjectures, which are regarded as dusty and outmoded whims; are also exposed to fashionable mockery. The ancients, namely, thought that three different types of life could be assumed to exist: *vegetative*, *animal* and *rational*. When the ancients combined the three immaterial principles of these three different types of life in man, they may well have erred. But when they distributed these immaterial principles among the three different classes of creature which grow and reproduce their kind, they were saying something which, although, of course, probably not capable of proof, was not for that reason absurd. (AA II: 330-331)

Here, Kant argues that we should not infer from the fact that an object lacks the undisputed external characteristic of life that it cannot be considered as living. His immediate concern is the undesired consequence that plants will turn out not to be living. He seems to argue that plants are probably alive, although their characteristic form of life, namely vegetation, is internal rather than external, and because of the way in which a plant is alive, it needs no external articulation necessary for voluntary or intentional action. Moreover, Kant reveals his awareness of the fact that this is rarely appreciated because the consideration of the difference between animal and vegetative life became unfashionable together with the philosophy that insisted on it: Aristotelian scholasticism. Against this fashion Kant defends the idea that considerations of the vegetative soul are legitimate, in spite of all other failures Aristotelianism might have, and that his contemporaries are misconstruing the issue of life when they are exclusively focused on the question whether animals have souls.<sup>9</sup> This is not to say that Kant embraces the notion of vegetative soul, or at least any more than that of animal soul, for he stresses that the recourse to immaterial principles is lazy philosophy, and that the mechanist approach is therefore the right way to go about investigating life, even though he is sympathetic to vitalism and animism for their recognition of the specificity of life. As a result, then, we find that Kant is unlikely to have regarded vegetative, unintentional teleology as counter to life.

When we turn to Kant's discussion, we see that the first kind of self-production he ascribes to natural purposes is that of reproduction:

First, a tree generates another tree in accordance with a known natural law.  
However, the tree that it generates is of the same species; and so it generates itself

---

<sup>9</sup> The question on whether animals have souls, or whether they are capable of sensation, of pleasure and pain and of consciousness, became extremely important in the wake of Descartes's infamous purported denial of animal life. Descartes's position is, however, best understood as a conveniently two-targeted attack on both the tendency of pyrrhonists since Montaigne to base skeptical arguments on the possibility of animal rationality, and the scholastic insistence that the explanation of animal behaviour and arguably even physiology requires recourse to a separate explanatory principle. The problem Descartes left may have been the result of the convergence and perhaps even conflation of the two debates. Descartes seems to have read Montaigne as collapsing the distinction between the animal soul and the rational soul by arguing that there is no clear line to be drawn between the sensitive, locomotive, and appetitive properties of animals and the rational properties of man. This would then imply that man has no ground to claim that he has a better access to knowledge than any other animal has (cf. Popkin 2003: 48-49). Descartes seems to believe as well that the rational and the sensitive properties are not fundamentally distinct, and therefore seeks to avoid Montaigne's conclusion by denying that animals have the latter properties at all. Thus, he arrives at a position where something is either fully rational or not alive at all. The aftermath of this problem in the 17<sup>th</sup> and 18<sup>th</sup> century reveals the difficulty of mechanism of not subscribing to this fundamental dualism between rational intentionality and dead necessity which ultimately, I surmise, underlies the more familiar dualisms of matter and soul.

as far as the **species** is concerned, in which it, on one side as effect, on the other as cause, unceasingly produces itself, and likewise, often producing itself, continuously preserves itself, as species. (AA V: 371)

One might wonder why Kant lists reproduction as a case of something being both cause and effect of itself, since it seems to be a relation between two (or three) different entities, where one brings about the other. This puzzle is best solved by reminding ourselves of some features of the traditional idea of reproduction as we saw it in the first section of this chapter, and of Kant's theory of epigenesis as I described it in the second chapter of this dissertation. In the first place, both Kant and Aristotle defend epigenesis because it ascribes to natural entities a genuine capacity to reproduce. Remember that the Platonic account in the *Timaeus* had serious difficulties in making reproduction a natural process, i.e. a process occurring in nature through the natural capacities of things. On this account, the creation of new biological individuals would require a new creative act either by the demiurg or by its world-immanent vicars. This is why Aristotle instead offered the theory of epigenesis, which states that biological individuals can produce their like without this production having to be one of intentional design. Similarly, Kant rejected preformation because it made all reproduction supernatural, and argued instead that we should attribute to natural things a capacity to reproduce, which he called a generative force. Thus, both Aristotle and Kant insisted that reproduction is a natural process occurring according to natural properties.

Nevertheless, both Aristotle and Kant also insisted that this reproductive capacity exhibits teleology because they did not believe material and mechanical processes could account for the specific form or structure of the end-product of generation. For this reason, they claim that the formative process is subject to further constraints, which specify what is otherwise underdetermined, i.e. "contingent", by material and mechanical processes. These specifying constraints are, according to both, species-constraints. The natural process of reproduction is thus one that is best understood as guided by the form or the idea of the species.

The two comments just made show that Kant could believe that reproduction involved natural teleology because he insisted on its *natural* rather than *supernatural* nature, and on its *form or idea-directedness* rather than *material-mechanical* nature. But this does not yet show why reproduction is a form of self-production. To understand this, we should take into account that Kant is here speaking of biological individuals insofar as they are of a specific species, i.e. of a certain kind rather than of another, which means that he is concerned here with the nature and properties of biological species. A reproductive process is one where the species is not just a classificatory concept, but a causally explanatory one: it explains why a certain kind of entity is produced through this process rather than another. Reproduction thus prompts us to commit ourselves to the existence of species as natural kinds. On the basis of this

assumption, Kant can characterize reproduction as the self-production of a species. In doing so, he adopts Aristotle's claim that reproduction is the means by which things perpetuate and produce themselves, not numerically, but as a kind. In fact, in *De Anima*, Aristotle already invoked this idea in order to show how both reproduction and nutrition are self-production:

[F]irst of all we must treat of nutrition and reproduction, for the nutritive soul is found along with all the others and is the most primitive and widely distributed power of soul, being indeed that one in virtue of which all are said to have life. The acts in which it manifests itself are reproduction and the use of food, because for any living thing that has reached its normal development and which is un mutilated, and whose mode of generation is not spontaneous, the most natural act is the production of another like itself, an animal producing an animal, a plant a plant, in order that, as far as its nature allows, it may partake in the eternal and divine. That is the goal towards which all things strive, that for the sake of which they do whatsoever their nature renders possible. The phrase 'for the sake of which' is ambiguous: it may mean either the end to achieve which, or the being in whose interest the act is done. Since then no living thing is able to partake in what is eternal and divine by uninterrupted continuance (for nothing perishable can for ever remain one and the same), it tries to achieve that end in the only way possible to it, and success is possible in varying degrees; so it remains not indeed as the self-same individual but continues its existence in something like itself-not numerically but specifically one. (415a-b)

It is rather obvious, therefore, that Kant is referring to this Aristotelian doctrine in the first example of the self-productive feature of the tree, and that he thereby embraces the idea that the consideration of species is important, and perhaps even indispensable, in the life sciences.

The second feature Kant lists is then unsurprisingly one of numerical or individual rather than specific self-production and self-perpetuation:

Second, a tree also generates itself as an **individual**. This sort of effect we call, of course, growth; but this is to be taken in such a way that it is entirely distinct from any other increase in magnitude in accordance with mechanical laws, and is to be regarded as equivalent, although under another name, with generation. This plant first prepares the matter that it adds to itself with a quality peculiar to its species, which could not be provided by the mechanism of nature outside of it, and develops itself further by means of material which, as far as its composition is concerned, is its own product. For although as far as the components that it receives from nature outside of itself are concerned, it must be regarded only as an educt, nevertheless in the separation and new composition of this raw material there is to be found an originality of the capacity for separation and formation in this sort of natural being that remains infinitely remote from all art when it attempts to reconstitute such a product of the vegetable kingdom from the



elements that it obtains by its decomposition or from the material that nature provides for its nourishment. (AA V: 371)

This passage is again reminiscent of the ancient theories of natural teleology espoused by Aristotle and Galen. First of all, Kant's claims that the process by which a biological individual produces itself is that of growth, which requires the power of nutrition, and is also a manifestation of the same power as that which is manifested, "under another name", in generation and reproduction, clearly refer to the Aristotelian doctrine that nutrition, growth and generation are properly viewed as collectively powers of the vegetative soul because they are in a way the same power. Additionally, Kant's claim that growth and nutrition cannot be understood by mere mechanical means echoes Aristotle and Galen as well, since both these thinkers stressed that nutrition involves three component capacities that, either singly or together, cannot be accounted for by the properties of material parts and their relative motions. The first power is that of attraction, whereby the nourished selects and draws in only the proper nutrients and ignores all others. The second is that of transformation, for living beings do not even use their proper nutrients in the form in which they receive them, but transform them into their own substance, tissue, etc. Finally, there is the process of intussusception, which we already encountered before (cf. 3.3.1), and which is responsible for the maintenance or production of the specific shape of the body or its parts. The reasoning behind postulating the latter power is that growth and nutrition cannot be understood as increase in magnitude through aggregation alone, but are better regarded as internal adaptation to a specific form.

Taken together, these comments reveal that Kant, together with Aristotle and Galen, regards physiology as revealing of the specificity of biological individuality. Biological entities must be understood as individuals which seek to maintain themselves by selectively attracting, transforming and assimilating materials that were initially alien to them. They cannot simply be understood as local mechanical and material processes that are not clearly delineated and individuated over and against their environments. In an important way, then, the life sciences do not just prompt us to invoke species as explanatory, but also to allow for individuating principles.

This clarification allows us to dispel a lingering confusion about Kant's concept of final causation. This confusion is explicit with Peter McLaughlin, who writes that "when Kant speaks of an idea or a concept as a 'purpose' or a 'final cause', he usually means the anticipation of the *product*, i.e. not the *causa finalis* in the proper sense but the *causa formalis*" (McLaughlin 1990: 39). I believe this is not a confusion on Kant's part, nor that it is peculiar, but that is rather revelatory of Kant's Aristotelianism on the topic. Kant is explicit about the fact that a natural purpose is to be considered as something which has its ends internal to it, which he usually phrases by stating that it *is* itself an end. This echoes Aristotle's views on natural substances, where the formal cause and the final

cause are supposed to coincide: the generative and nutritive processes in living beings are properly understood only insofar as they are understood as for the sake of the production and perpetuation of the species and the individual.

We can now also better understand what Kant means when he says that natural purposes too prompt us to regard them as possible only according to a certain concept. This needn't mean that we can understand them only as intended to have this form, because it can also mean that these natural purposes need to be understood by taking recourse to the concept of their species or of their individual nature.<sup>10</sup> As Kant states elsewhere in the "Analytic of the Teleological Power of Judgment", what we need to understand about a natural purpose is its unity according to a certain rule (AA V: 360). I suggest we understand this as meaning that the processes we characterize as naturally teleological prompt us to regard them as if they were coordinated in order to accord with a certain norm, i.e. that they are best understood as exhibiting discriminating properties which suggest that they are guided by a normative rather than a factual idea, namely that of *having to preserve* a given structure or entity. The question then, of course, becomes that of how natural processes can be norm-governed despite being non-intentional. This is precisely the puzzle of the concept of a natural purpose, the puzzle that qualifies the concept as incomprehensible.

It is important, however, that we discuss the final trait of a natural purpose that Kant lists in his example of a tree, since it has important consequences for our understanding of what Kant means by the previous two. This trait is best described as that of *plasticity* (Cf. Zuckert 2007: 99):

**Third**, one part of this creature also generates itself in such a way that the preservation of the one is reciprocally dependent on the preservation of the other. An eye from the leaf of one tree grafted into the twig of another brings forth a growth of its own kind in an alien stock, and similarly a scion attached to another trunk. Hence one can regard every twig or leaf of one tree as merely grafted or inoculated into it, hence as a tree existing in itself, which only depends on the other and nourishes itself parasitically. At the same time, the leaves are certainly products of the tree, yet they preserve it in turn, for repeated defoliation

---

<sup>10</sup> Rachel Zuckert has remarked that "[t]hough Kant initially presents the three aspects of organisms as equally important, he tends to neglect the first in his further discussion" (Zuckert 2007: 99). I fear this is a bit too precipitous, because on my reading the discussion of species identity remains constantly in view insofar as the individual is best understood not only as this specific thing (*tode ti*, in its haecceity), but also as this specific kind of thing (in its quiddity). For Aristotle, the latter always contributed importantly to the former, since the kind-terms constitutes the sortal which allows the individual to be numerically identified. Of course, Kant is, as Zuckert rightly remarks, more concerned here with individuality than with species persistence, but he may have taken himself to have written on the latter subject *ad nauseam* already in his various texts on epigenesis.

would kill it, and its growth depends upon their effect on the stem. The self-help of nature in the case of injury in these creatures, where the lack of a part that is necessary for the preservation of the neighboring parts can be made good by the others; the miscarriages or malformations in growth, where certain parts form themselves in an entirely new way because of chance defects or obstacles, in order to preserve that which exists and bring forth an anomalous creature: these I mention only in passing, although they belong among the most wonderful properties of organized creatures. (AA V: 371-372)

This passage actually lists two forms of plasticity that are best discussed separately. The first form is that whereby a part of a natural purpose, when removed, persists in its function as a part despite removal from the whole in which it is embedded. Kant gives the example of grafting, whereby a transplanted twig retains its specificity as a part of the previous tree, even though it now flourishes in another tree that may be of a different species. This suggests to Kant that the parts of natural purposes do have, to some extent, an individuality and persistence of their own apart from their parent whole, and that these parts can therefore be viewed as individuals in their own right. This is expressed in his comment that the branches and twigs of trees can be seen as distinct biological individuals parasitic upon the supposed whole in which they figure. But he equally notes that this is misleading as a characterization, because these parts are not merely parasitic on the whole or on other parts, for they serve to preserve the others. Kant thus describes an organism here as a kind of coordinated symbiosis rather than as an unequivocal individual or unity.

The second form of plasticity is that of compensation, whereby upon injury or damage of one part the other parts can recoordinate such as to compensate for the functional loss. This resilience of biological individuals is very peculiar, for it seems to run counter to the idea of an individual essence. In discussing the second trait of a natural purpose, Kant stressed that the concept of an individual is central to our understanding of biological processes. This concept is supposed to be the criterion according to which a structure counts as that individual and no other. In its turn, the criterion seems to be a specific form or structure, a relation between certain parts. But now Kant seems to be saying that an individual can change the proportion and relation between its parts and the specific nature of those parts *in order to* persist as a functional and functioning whole. At first glance, the idea of an individual essence would imply that the whole is no longer functioning, or malfunctioning, and thus either deceased or diseased when its criterion or individuation is no longer strictly observed and conserved.

I would like to submit that the solution to this problem lies in the fact that, in his consideration of biological individuality, Kant's position is akin to that of Georges Canguilhem:

the pathological or abnormal state does not consist in the absence of every norm. Disease is still a norm of life but it is an inferior norm in the sense that it tolerates no deviation from the conditions in which it is valid, incapable as it is of changing itself into another norm. The sick living being is normalized in well-defined conditions of existence and has lost his normative capacity, the capacity to establish other norms. (Canguilhem 1991: 183)

In this passage, Canguilhem is suggesting that the persistence of a biological individual is not conditional on its maintaining a certain norm, but rather on its capacity to change this norm. In the event of damage, or of changing outside conditions, a healthy and functioning organism can change its specific structure to respond to these alterations and compensate for losses. A diseased individual is precisely one which is incapable of such a renormalization, of such a shift in norms: it is one whose norm, whose concept, whose essence is no longer plastic, and thus lacks responsiveness.

That Kant regarded biological essences as essentially plastic is unsurprising in the light of his theory of epigenesis. After all, as we saw, the constraints on the generative force that form the specific constraints can change over time, and are thus plastic, although not indefinitely so. Moreover, Kant stated in his discussion of the second form of self-production that the power of reproduction and the power of nutrition and growth are different manifestations of the same power. As a result, we can expect that the latter exhibits the same plasticity as the former. Plasticity and adaptability, then, are built into the Kantian conceptions of biological species and individual.

One final remark that I would like to make before leaving this discussion of what are ultimately Kant's preliminaries to his proper discussion of the concept of natural purpose, is that the concept of plasticity reveals why Kant would not be tempted to think feedback and equilibration mechanisms could account for the natural processes he listed in §64. The reason for this is that the stability of an organism is never actually one of equilibrium, for if so it would lack the plasticity and responsiveness which Kant deems somehow essential to life. An organism never fully actualizes its form, but always maintains a certain room for potential alternative ways of being. The stability of the organism on such a view would best be characterized, in the terms of Gilbert Simondon (1989: 18-19), as a metastability, a stability away from equilibrium. Mechanicist conceptions of equilibrium, of balance, of feedback, are at place only in an entity which is always already fully formed and no longer capable of change and responsiveness – a preformationist conception of the organism.

In listing all these peculiar properties of natural purposes, Kant himself realizes he has made it all the more urgent to explain how one could even consider such an entity. After all, these preliminaries all point in the same direction, namely the idea of a natural process which involves concepts of reason as causally explanatory without involving intentionality and volition. Kant now has to face the task of clarifying the concept of such a process or entity.

### 4.2.3 Natural Purposiveness Determined

The central paragraph of the “Analytic of the Teleological Power of Judgment” is §65, where Kant proposes to derive the concept of something which is reciprocally cause and effect of itself, which he admits to be “somewhat improper and indeterminate” (AA V: 372), from a more determinate concept. Such a derivation would require, of course, that one determine what it means to be cause and effect of oneself, which in turn requires a gloss on the possible constraints on the relation of cause and effect. It is unsurprising, then, that Kant starts off his treatment with a discussion of causality, or rather, with the different commonly accepted forms of causality. The first one is that of efficient causes:

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; and the things themselves, which as effects presuppose others as their causes, cannot conversely be the causes of these at the same time. This causal nexus is called that of efficient causes (*nexus effectivus*). (AA V: 372)

This first form of causality seems to be that of mechanical causation, which is strictly linear, and proceeds from cause to effect. This is borne out by Kant’s use of the term “descending”, by which he usually means a progression in a series from the anterior to the posterior, or from the antecedent to the consequent. Nevertheless, this characterization is somewhat misleading, because there is a sense in which mechanical causation can be reciprocal, namely in the sense of a mechanical system or equilibrium. The solar system, for Kant, can be regarded as mechanical system, even though the attractive relations work both ways, and the movement of every planet or planetoid is determined by the movements of all others. In this case, however, the movements of the objects in the systems can be understood as composed of all the simple movements they would have if they would only be causally related to one other object of the system. Efficient causation therefore does not imply that a system be understood as if all relations of cause and effect in it are linear and descending, but only that all causal relations in the system can be fully analyzed into causal relations that are linear and descending. Presumably, the reciprocal causality of natural purposes would not be of this kind.

Another issue with this passage is that of its scope. It seems as if Kant is saying that the understanding must think of causal relations as descending. This raises the question as to whether Kant thinks this restriction is imposed by the category of causality or not. At least on first sight, it seems as if the linearity is required by the category of causality, since Kant calls the latter “the concept of the relation of cause and effect, the former of which determines the latter in time, as its consequence” (CPR A 188 / B 234). But then it would seem that mechanical causation is required by the very category of causality, which would mean that no sense can be made of the idea of natural teleology. As will

become clear in the rest of my discussion, I believe it is right to read Kant as stating that reciprocal causation violates mechanicism, but that it is not right to claim that mechanicism is required by causality in general. I believe the reason for this is that mechanicism is a claim about the rules that connect causes and effects, and the way in which these rules relate to each other, rather than a claim about the direction of causal relations.

In spite of this important restriction, Kant believes that we can grasp the idea of a causal connection where the relation between cause and effect is not linear and descending, for he writes that

a causal nexus can also be conceived in accordance with a concept of reason (of ends), which, if considered as a series, would carry with it descending as well as ascending dependency, in which the thing which is on the one hand designated as an effect nevertheless deserves, in ascent, the name of a cause of the same thing of which it is the effect. (AA V: 372)

This passage is puzzling, because Kant seems to be saying that we *can* conceive of a form of causality where the final cause is as much a cause as the efficient cause is. But this would violate the temporal restriction required by the category of causality. Peter McLaughlin (1990: 48-49) explicitly signaled this problem, and argued that, because “[a] consideration of some kind of ‘backwards causality’ is strictly excluded by the Kantian concept of causality”, “[t]he interpretation that might suggest itself of Kant’s metaphor of the ‘series’ of cause and effect as forwards or backwards in time is untenable”. For this reason, he suggests we take Kant to mean not the strict temporal causal relation between cause and effect, but rather the simultaneous and spatial mereological relation between part and whole.

I believe this move is mistaken, not just because it requires that we read past Kant’s own explicit wording (which may sometimes be necessary), but also because, as we will see, Kant did not believe such a causal relation involved contradiction, even though he would have to hold it to be incomprehensible. But this is not surprising, since in the previous paragraph Kant stated explicitly that the form of purposiveness under consideration, namely natural purposiveness, is non-contradictory but incomprehensible. I therefore suggest we take Kant to really mean that this concept of causality is thinkable but not comprehensible, and try to make sense of that claim.<sup>11</sup>

---

<sup>11</sup> Unlike many other commentators, Rachel Zuckert does not seem to have a problem accepting that the reciprocal relation invoked by Kant as that of natural purposiveness is in fact a temporal one (Cf. Zuckert 2007: 117-118). This is because she stresses the plasticity of organisms, as well as their capacity for anticipation. In an early paper of mine (Demarest 2013b), I likewise argued that Kant’s natural purposes are best understood as anticipatory systems.

After introducing this peculiar kind of causality, Kant states that there is a form in which it is not really that objectionable at all:

In the practical sphere (namely, of art) such a connection can readily be found, e.g., the house is certainly the cause of the sums that are taken in as rent, while conversely the representation of this possible income was the cause of the construction of the house. Such a causal connection is called that of final causes (*nexus finalis*). (AA V: 372)

I take Kant to say here that we are familiar with a kind of ascending causality in the case of intentional and volitional action. This is the form of causality usually referred to as final causation. In order to understand that Kant is here referring to the classical, scholastic version of the notion of a final cause, we need to note that he describes the end as a certain benefit to be gotten from the existence of that which is brought about, namely the rent. This allows us to read the relation between efficient and final causality here as akin to that suggested by Aquinas (2014: 9) with his phrase “an end is the cause of the causality of the efficient, because it brings it about that efficient is efficient”. In one sense, the efficient cause is in fact the cause of the end, but in another, this efficient cause would not have operated without the end. That which makes an efficient cause efficient, or brings it into motion, is best understood as a “motive”, at least in the case of intentional action. This brings us to the traditional scholastic notion of a final cause, which required that the final cause be temporally prior to the end of a cause in the form of an intention or representation in the mind of the one undertaking the efficient causation. The idea of intentionality, then, is used to reconcile the notion of final causation with the temporal restriction, but on pains of equivocating between the concept of cause, as Kant seems to realize:”The first could perhaps more aptly be called the connection of real causes, and the second that of ideal ones, since with this terminology it would immediately be grasped that there cannot be more than these two kinds of causality.” (AA V: 372-373) Kant suggests that, on this explanation, the distinction is not one between efficient and final, or between descending and ascending causation, but between two kinds of cause. The one is the cause in the sense of the real processes taking place in the world, whereas the latter is the cause in the sense of the intentions and motives, which are ideas, of the free agent undertaking the construction. Moreover, Kant clarifies that according to this distinction only these two kinds of causality are possible, namely mechanical causation and free intentional action.

In the literature, there has been a marked tendency to believe that Kant himself admitted these to be the only possible kinds of causation. I will argue, however, that this is wrong, and that they are merely the only concepts of causation comprehensible to us as finite cognizers. Before doing so, however, I will briefly discuss some issues with trying to fit Kant’s treatment of natural purposiveness within the confines of this dichotomy or antinomy. Such a fit would require that we understand organisms as

somehow suggesting the causality of freedom. The most obvious way to do this is to read Kant as saying that organisms exhibit the causation of freedom because they exhibit *intentional action*. We have already seen, and will see again, that this option is unsatisfactory because Kant seems constantly concerned with showing that organisms cannot be understood as designed.

But if not intentionality, then which trait of freedom do natural purposes exhibit? Clark Zumbach has offered the suggestion that it is spontaneity to which we must turn: “One way of understanding Kant’s position is to view it as the claim that we must view living processes in terms of the idea of randomness; that is, we might consider these occurrences as inherently (not just descriptively) random occurrences” (Zumbach 1984: 103). This proposal has an initial attractiveness, since it would do justice to the plastic and adaptive traits Zumbach admits Kant’s “epigenetic systems” to have. But it does not withstand further scrutiny, because Zumbach is merely pushing onto Kant the dogmas of the philosophical spokesmen of mainstream biology in the mid and late 20<sup>th</sup> century, like Ernst Mayr and Jacques Monod. For Mayr (1997) and Monod (1970), the specificity of the biological subject matter was due to its being produced by both necessity and contingency. In the forms of mutation and of variation within a biopopulation, a random and non-deterministic element creeps into evolution and genetics.

But this is mere appropriation, inspired by the cultural dominance of the modern evolutionary synthesis and molecular biology at the time of writing of Zumbach’s monograph. Zumbach is here desperately, and against textual evidence, trying to read Kant as embracing the “sensible” mainstream biology of the 20<sup>th</sup> century rather than the heterodox perspectives that would again resurface after the publication of Zumbach’s book, with the rise of systemic approaches, with the reintroduction of epigenetics and with the renewed attention to the recalcitrant fields of developmental biology, ecology and ethology. Kant’s own concept of a natural purpose fits more with these heterodoxies than with the mechanistically inspired modern synthesis. Moreover, Kant’s claim that the concept of natural purpose and its peculiar features are indispensable and ineradicable in the life sciences seem vindicated by the recurrent rise of heterodoxies and sentiments of failure regarding the 20<sup>th</sup> century’s mechanistic paradigm in so much research in the life sciences.

I would like to argue, however, that there is one feature of freedom that is more clearly exhibited by natural purposes, in spite of their being natural, and this is *normativity*. It is sometimes too readily forgotten that Kant’s talk of intentionality and of spontaneity in the case of humanity is there mostly to make room for the aspect of normativity, i.e. of the *ought* over and against the *is*. This is clear from the manner in which Kant characterizes freedom, namely as autonomy of reason. By this he does not



mean that human action is undetermined, but rather that we need to think of it as self-determined.<sup>12</sup> Spontaneity, on such a reading, is not the capacity for random or fickle action, but rather the capacity to be determined not (only) by causes, but (also) by reasons, reasons which have a normative character in that they determine how one ought to behave. This does not just mean moral obligation, but also *epistemic* obligation. Knowledge, for Kant, is more a matter of what one ought to accept, what one ought to think, how one ought to justify, than a matter of what is and how one acquires access to it.<sup>13</sup> In many ways, transcendental idealism is the claim that our ontological commitments are dependent upon our epistemic commitments, and our epistemic commitments are normative commitments just as much as our moral commitments are – hence the unity of the theoretical and the practical. Given that, for Kant, freedom means autonomy, and autonomy means determination by justification alone, a random event would not be free precisely because it lacks justification. Similarly, human intentionality is understood as the capacity to set, picture and recognize ends, whereas spontaneity is the capacity to act in accordance with these ends rather than in accordance with causes.

In the previous subsection, I already argued that an organism is best understood as exhibiting normativity, in the sense that it is not obviously mechanistically determined to maintain its form, but rather that it seems to strive to instantiate and preserve a

---

<sup>12</sup> Cf. AA V: 33, where Kant contrasts a negative and a positive sense of freedom, the negative being the fact that reason, in morality, is not determined by external ends set by nature, and the positive being the fact that reason can act according to its own principle, which yields obligation.

<sup>13</sup> The idea that, with Kant, we see a decisive shift to epistemology as dealing with a normative enterprise first became explicit in the work of Wilfrid Sellars, and was subsequently articulated by Sellarsians like Robert Brandom, who wrote: “The nature and significance of the sea change from Cartesian certainty to Kantian necessity will be misunderstood unless it is kept in mind that by 'necessary' Kant means 'in accord with a rule'. It is in this sense that he is entitled to talk about the *natural* necessity whose recognition is implicit in cognitive or theoretical activity, and the *moral* necessity whose recognition is implicit in practical activity, as species of one genus. The key concept of each is obligation by a rule. It is tempting, but misleading, to understand Kant's use of the notion of necessity anachronistically, in terms of contemporary discussions of alethic modality. It is misleading because Kant's concerns are at base normative, in the sense that the fundamental categories are those of deontic modality, of commitment and entitlement, rather than of alethic modality, of necessity and possibility as those terms are used today. Kant's commitment to the primacy of the practical consists in seeing both theoretical and practical consciousness, cognitive and conative activity, in these ultimately normative terms.” (Brandom 1994: 10) This realization is important, because it marks the major distinction between Kant's transcendental idealism and phenomenalism or idealism: the latter views are all committed to the fact that reality ultimately consists more of the mental or intentional entities that we are more directly acquainted with. Transcendental idealism, however, is the view that such debates on what there is are futile and ultimately even harmful. Instead, it replaces talk of “is” and “reality” with talk of “ought” and “obligation”. To express it in the form of a boutade, with the obvious qualification that it is but a boutade: for transcendental idealism, all that there “is” is “oughts”.

certain form and nature. A natural purpose seems as if it is committed to, as if it *ought* to instantiate this form, this form being its proper and internal end. The following passage from the first, unpublished Introduction to the *Critique of the Power of Judgment* explicitly phrases the specificity of natural purposes in terms of “oughts”:

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). That we can see with the eye we experience immediately, as we do the outer and internal structure of the eye, which contain the conditions for its possible use, and therefore its causality in accordance with mechanical laws. But I can also use a stone, either in order to crush something upon it, or to build something upon it, etc., and these effects can also be related to their causes as ends, although I cannot on that account say that it ought to have served for building. Only of the eye do I judge that it ought to have been suitable for seeing, and although its figure, the character of all its parts and their composition, judged in accordance with merely mechanical laws of nature, is entirely contingent for my power of judgment, I nevertheless think in its form and in its construction a necessity for being formed in a certain way, namely in accordance with a concept that precedes the formative causes of this organ, without which the possibility of this product of nature is not comprehensible for me in accordance with any mechanical natural law (which is not the case with the stone). Now this ought contains a necessity which is clearly distinguished from physical-mechanical necessity, in accordance with which a thing is possible in accordance with mere laws of efficient causes (without any preceding idea of that thing), and can no more be determined through merely physical (empirical) laws than the necessity of the aesthetic judgment can be determined through psychological ones, but instead requires its own a priori principle in the power of judgment, insofar as it is reflecting, under which the teleological judgment stands and by means of which both its validity and its limitation must also be determined. (AA XX: 240-241)

If this is how Kant interpreted natural purpose, then this would explain why he saw it as mechanically inexplicable, and as suggesting an analogy with human freedom, for we only know such normative behavior in the form in which we exhibit it. We now what it is to have internal ends, to be determined by these internal ends, and to be committed to these ends, in the peculiar feeling of freedom encountered in our obtrusive moral considerations and self-incriminations. But human freedom has the peculiarity that it suggests that we take ourselves, insofar as we are moral and epistemic agents, as somehow (not obviously) natural. In the case of *natural* purposes, however, we can do no

such thing, for in their case we need to make sense of them as somehow natural and normative, i.e. naturally normative.

Having made these comments on the concept of causality, Kant continues with a further characterization of the concept of a natural purpose, listing two criteria. The first criterion is what I will call the holism-criterion:

Now for a thing as a natural end it is requisite, first, that its parts (as far as their existence and their form are concerned) are possible only through their relation to the whole. For the thing itself is an end, and is thus comprehended under a concept or an idea that must determine a priori everything that is to be contained in it. But insofar as a thing is conceived of as possible only in this way it is merely a work of art, i.e., the product of a rational cause distinct from the matter (the parts), the causality of which (in the production and combination of the parts) is determined through its idea of a whole that is thereby possible (thus not through nature outside of it). (AA V: 373)

First of all, Kant seems to clarify here what he meant when he said that, when we consider a thing as an end, we are comprehending it under a concept or idea that must determine a priori everything that is to be contained in it. The clarification adds that to understand a thing in this way is to regard its parts as possible only through their relation to the whole. This is in itself unsurprising, given the fact that organisms have throughout history prompted holistic conceptions, and given the specific features of natural purposes Kant listed in his discussion of the example of a tree. The parts of a natural purpose are to be understood as together constituting an individual of a certain kind (species), in such a way that the coordination of these parts is what allows the natural purpose to be this individual of this kind (species).

In spite of the familiarity of this holist criterion to organismic thinking, Kant goes on to warn us that it is in itself insufficient and misleading, because there are things that exhibit such a holist structure without being natural purposes, namely artifacts and machines. In the case of an artifact, in its literal sense of “work of art”, the whole does determine the nature and relation of the parts, because these parts have been selected and disposed by a rational agent in such a way that they together bring about the desired effect. What Kant finds problematic here is that the whole, insofar as it is productive of the parts and their relations, is not internal to the natural purpose itself, but external to it, as the representation, idea or concept in the rational artificer. This is why Kant offers a second criterion for being a natural end:

But if a thing, as a natural product, is nevertheless to contain in itself and its internal possibility a relation to ends, i.e., is to be possible only as a natural end and without the causality of the concepts of a rational being outside of it, then it is required, second, that its parts be combined into a whole by being reciprocally the cause and effect of their form. For in this way alone is it possible in turn for the

idea of the whole conversely (reciprocally) to determine the form and combination of all the parts: not as a cause – for then it would be a product of art – but as a ground for the cognition of the systematic unity of the form and the combination of all of the manifold that is contained in the given material for someone who judges it. (AA V: 373)

I will call this second criterion the naturalness criterion, for it is the criterion for being not just an end, not just purposive, but a *natural purpose*, i.e. *naturally purposive*. Such a natural purpose should be purposive “without the causality of the concepts of a rational being outside of it”. Here, Kant is explicit that the concept of a natural purpose does not suggest, but rather precludes, intentional agency.<sup>14</sup> It is clear that he is talking about non-intentional, natural teleology, the conception of teleology that was central to Aristotle’s theory. Now, just as he clarified the idea of purposiveness in the first criterion through the determination of the parts by the whole, here he clarifies the idea of natural purposiveness as a reciprocal determination of the parts. Understanding this puzzling, and at first sight inconsistent characterization of a natural purpose will allow us to understand the revolutionary nature of Kant’s concept of an organism.

The major peculiarity about Kant’s treatment here is that it doesn’t just add another criterion to the holist criterion, but rather that he *qualifies* the first criterion. In fact, Kant states that in the case of a natural purpose the whole is not a cause of the nature and relation of the parts in the sense that it is prior to and distinct from this nature and relation of the parts. The judgment of holism is not prompted by there (necessarily) being some or other whole that is in some sense prior to the parts, but rather by the fact that the different parts are not causally prior to one another in spite of their intimate and intricate interconnection. The nature of each part is in a way due to its relation to other parts, and its relation to other parts due to the respective natures of all involved parts. This radically changes the sense Kant attaches to circular causation: it is not the coexistence of a descending causal relation from part to whole and an ascending causal relation from whole to part, but rather the coexistence of a large amount of reciprocal causal relations between parts and other parts. Kant’s picture of the organism, then, is not that of a harmonious whole due to the imposition of a general form or structure, but rather that of a harmonious whole due to the reciprocal formative actions of the parts:

---

<sup>14</sup> We can see why this is so by reflecting on the following passage from Dante Alighieri’s *De Monarchia*: “according to the intention of the creator, as creator, the ultimate end of a created being is not the being itself but its proper function.” (1904: 10) Dante here expresses the creationist view that all created entities exist for the sake of something, but that they can never exist for the sake of themselves. Kant’s concept of a natural purpose is precisely that of an end in itself. The idea of intentional creation thus does not form a good framework for understanding natural purposiveness.

For a body, therefore, which is to be judged as a natural end in itself and in accordance with its internal possibility, it is required that its parts reciprocally produce each other, as far as both their form and their combination is concerned, and thus produce a whole out of their own causality, the concept of which, conversely, is in turn the cause (in a being that would possess the causality according to concepts appropriate for such a product) of it in accordance with a principle; consequently the connection of **efficient causes** could at the same time be judged as an **effect through final causes**. (AA V: 373)

In natural purposes, it is the causality of the parts, insofar as these mutually determine each other, that are responsible for the whole, and the idea or concept of the whole is due to the peculiarly coordinated, concerted co-production of these parts. Such parts are, Kant specifies, *organs*:

In such a product of nature 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, i.e., as an instrument (organ), which is, however, not sufficient (for it could also be an instrument of art, and thus represented as possible at all only as an end); rather it must be thought of as an organ that **produces** the other parts (consequently each produces the others reciprocally), which cannot be the case in any instrument of art, but only of nature, which provides all the matter for instruments (even those of art): only then and on that account can such a product, as an **organized** and **self-organizing** being, be called a **natural end**. (AA V: 373-374)

For most of its history, the word organ maintained its original meaning of an instrument in general, not just an organ of an organism. The reason for assimilating the latter to the former is that the organ is conceived as an instrument of which the organism makes use in its specific way of life. For Kant, however, these two senses of organ cannot so easily be assimilated to each other on pain of equivocation: an organ, in the sense of an instrument of life, does not merely coordinate its actions with other organs, nor is it able to contribute merely because of the natures of other organs; it *produces* other organs (usually with the assistance of yet other organs) and is in turn *produced* by those organs (usually with the assistance of yet other organs). An organ, on this view, is a product of its products, is productive only due to the productivity of its products.

It is on this feature that Kant focuses in order to distinguish organisms from artifacts, amongst others by discussing the example of the watch, a favourite example of mechanistic natural theology:

In a watch one part is the instrument for the motion of another, but one wheel is not the efficient cause for the production of the other: one part is certainly present for the sake of the other but not because of it. Hence the producing cause of the watch and its form is not contained in the nature (of this matter), but

outside of it, in a being that can act in accordance with an idea of a whole that is possible through its causality. Thus one wheel in the watch does not produce the other, and even less does one watch produce another, using for that purpose other matter (organizing it); hence it also cannot by itself replace parts that have been taken from it, or make good defects in its original construction by the addition of other parts, or somehow repair itself when it has fallen into disorder: all of which, by contrast, we can expect from organized nature. – An organized being is thus not a mere machine, for that has only a **motive** power, while the organized being possesses in itself a **formative** power, and indeed one that it communicates to the matter, which does not have it (it organizes the latter): thus it has a self-propagating formative power, which cannot be explained through the capacity for movement alone (that is, mechanism). (AA V: 374)

I believe Kant is saying here that, in an artifact, the typical circular causality obtaining between parts in organisms is not present, for the way that parts relate to each other insofar as their form and their existence is concerned is not a causal one. Admittedly, the parts of an artifact interact with others, and its causal behaviour was intended to be constrained in a particular way by the parts to which it relates, but it is not itself the cause of these constraints. In an organism, each part is to be considered as constrained due to parts that are in a way due to its specific causal behaviour. In this manner, the causal properties of the parts of an organism are self-determining and self-constraining. The difference thus becomes the following: the parts of an organism can behave/move in any of a number of ways, but their being surrounded by specific other parts ensures that they only behave/move in a very (limited number of) specific manner(s). These specific movements are, then, contingent to the motive force of these parts. The parts of an organism, however, do not only have the capacity to behave/move in any number of ways, they also have the capacity to produce/form the surrounding part that constrain its own capacity to behave/move. For an organic part, its specific behaviour is not contingent to it, but essential, even though it seems to be contingent to its motive properties alone.

Kant does not, however, think we can readily understand this property of living beings, or the species of causality it suggests. In fact, he stresses that it is “inscrutable” and “not analogous with any causality that we know”:

One says far too little about nature and its capacity in organized products if one calls this an **analogue of art**: for in that case one conceives of the artist (a rational being) outside of it. Rather, it organizes itself, and in every species of its organized products, of course in accordance with some example in the whole, but also with appropriate deviations, which are required in the circumstances for selfpreservation. Perhaps one comes closer to this inscrutable property if one calls it an **analogue of life**: but then one must either endow matter as mere matter with a property (hylozoism) that contradicts its essence, or else associate with it an

alien principle **standing in communion** with it (a soul), in which case, however, if such a product is to be a product of nature, organized matter as an instrument of that soul is already presupposed, and thus makes that product not the least more comprehensible, or else the soul is made into an artificer of this structure, and the product must be withdrawn from (corporeal) nature. Strictly speaking, the organization of nature is therefore not analogous with any causality that we know. (AA V: 374-375)

In this passage, Kant first points out that we have not understood living beings at all if we merely attribute to them a causality analogous to practical causality, i.e. the causality of art. In the case of art, the whole, in the form of the idea, is external to the product, whereas in the case of life, the artificer is the artifact itself. But he goes on to dismiss both hylozoist and animist ways of doing justice to this property. The Hylozoist explains the self-organizing property by ascribing this property to the matter of the organism, whereas the animist explains it by postulating a non-material principle with intentional or quasi-intentional capacities that is somehow responsible for the form and the formation of the organism. For Kant, hylozoism is problematic, because, for him, the notion of matter means that which is causally passive and inert<sup>15</sup>, and hence no sense can be made of the idea of active matter. Animism, on the other hand, must either place the guiding principle in the matter of the organism itself, in which case it collapses into hylozoism, or place it outside (not necessarily spatially, of course) of the organism, in which case it collapses into the view of self-organization as analogous the practical causality.

Before going on to discuss some further important consequences of Kant's idiosyncratic conception of the organism, I would like to tackle a recurrent challenge in the literature. This challenge surfaces best in Hannah Ginsborg's claim that there are two kinds of mechanical inexplicability in Kant: "the first supporting a teleological conception of organisms, and emphasizing the analogy between organisms and artifacts; the second supporting a conception of organisms as natural, and emphasizing the way in which organisms differ from artifacts." (Ginsborg 2004: 60) These two kinds of mechanical inexplicability are of course intended to correspond to the two criteria for natural purposiveness, the first being a criterion for being an end, the second one for being a natural end.

---

<sup>15</sup> This is not merely a stipulation of Kant's of course. His reasoning seems to be that the notion of matter plays a specific role in natural philosophy by its difference from certain other concepts. A very important concept for Kant is of course that of a *force*. A force is not a metaphysical entity, on Kant's account, but a physical one which means "that which accounts for motion". On such a reading, Kant would be able to conclude that the notion of active matter is nonsensical because it would mean that matter is a force, and a force is supposed to be that operating *on* matter.

Although there is some truth to Ginsborg's analysis, she, like many other authors, may have misconstrued the sense in which artifacts are supposed to be mechanically inexplicable. Recall the discussion of the relation between mechanicism and teleology in 4.1.2. There, we saw that the mechanistic philosophy is in no way inhospitable to the notion of teleology, as long as the teleology is adequately understood. For mechanists, the only admissible explanations are those in terms of the general principles of movements insofar as they work in specifically spatially configured systems. Thus, two elements are required: general principles of motion and specific structures. In themselves, the general principles of motion underdetermine their outcomes in these systems, for the determination equally comes from the specific spatial configurations, or, in later systems, the specific values entered into the argument-positions of the functions expressing these general principles. Now, for mechanists, the only powers that nature has are the general principles of motion, which by themselves cannot account for the structure of nature, for this is partly due to some initial spatial configuration, some initial input. This raises the question of whence this initial structure is derived, if it is not to be a brute fact about nature. The early mechanists answer this question by stating that *God* has determined the initial structures or inputs. The debate is then on whether the reasons of *God's* choice for these structures and values are accessible to us or not (Descartes said no, Boyle said yes).

There is therefore something very peculiar about saying that artifacts are mechanically inexplicable because their specific structure cannot be explained by the general principles of mechanics, because in this sense, *all* mechanical systems are mechanically inexplicable. This is the reason why, for mechanists, the structure of the universe suggests the hand of *God*, for any other range of structures or inputs might have led to a much less satisfying (in whatever sense they may attach to this word) universe. The interesting feature of the mechanistic paradigm is that it can neatly distinguish between motion and structure, because it assumes they correspond with different kinds of causality. In this way, mechanists have found it useful to regard *every* physical system as if it were an artifact. But then there is no class of entities which specifically suggests design in nature, as long as we take it to be so that all, or at least most, alternative configurations would have had less interesting results.

I would like to advance, therefore, that neither for Kant nor for any of his contemporaries, there is something interestingly mechanically inexplicable about artifacts. The problem arises only if we regard the structure and the form of these artifacts as due to *nature*. What emerges from this distinction is the strained relationship between naturalism and mechanicism in the sciences, where naturalism is understood as the claim that everything in nature is to be understood through nature, and mechanicism as the claim that everything is due to the most general of principles of motion. If nature only acts generally, then it obviously cannot account for specificity, which results in the defender of mechanicism having to invoke a supernatural principle



in order to account for specificity, thereby violating his supposed naturalism. I submit that, in his discussion of the notion of a natural purpose, Kant was attempting to formulate the idea that artifacts are mechanically explicable at the expense of being naturally inexplicable, whereas self-organizing systems are considered to be naturally explicable at the expense of being mechanically inexplicable. This is the puzzle that will give rise to the antinomy of the teleological power of judgment, as we will see in the next section.

Before passing on to the antinomy, however, I would like to briefly dispel two possible misgivings about Kant's conception of a natural purpose. The first is that of his much decried panfunctionalism, and the second is the solution to the issue of whether ecologies can be natural purposes, which we left hanging at the end of 4.2.1.

After concluding his presentation of the concept of a natural purpose in §65, Kant offers, in §66, a principle for the judging of the internal purposiveness in organized beings: "This principle, or its definition, states: **An organized product of nature is that in which everything is an end and reciprocally a means as well.** Nothing in it is in vain, purposeless, or to be ascribed to a blind mechanism of nature." (AA V: 376) This principle has met with serious skepticism, because its assumption that in a natural purpose everything is functional seems to be outrageously implausible, if not empirically false. Our reticence towards this panfunctionalism in Kant may, however, be due to our different interpretation of the notion of function. In order to appreciate this, consider one of the most famous and effective critiques of panfunctionalism in evolutionary biology, offered by Stephen J. Gould and Richard Lewontin. Gould and Lewontin (1979) criticized what they called the adaptationist program for ultimately coinciding with the Panglossian paradigm. They mean by this the views of the character doctor Pangloss from Voltaire's *Candide*, who defended Leibnizian optimism, i.e. the view that this is the best world and that everything is for the best, through laughable explanations of for instance the nose as something that is there for supporting glasses. For Gould and Lewontin, the adaptationist program is Panglossian because it treats every part of the organism as an individual trait, which is then thought to be present because it makes an individual contribution to the optimization of the fitness of the organism, or because its presence together with other traits leads to a Pareto-optimal<sup>16</sup> situation in the organism (where, of course, no part can be more optimal without decreasing the optimality of other parts). Against this, they argue that there are in organisms, much like in designed structures such as architectural products, elements that do not

---

<sup>16</sup> It is not difficult to see in Gould and Lewontin's two steps of studies in the adaptationist programme as that of first determining the optimality of individual traits, and then determining the Pareto-optimality of a group of traits. I believe it is not too far-fetched to suggest that at least Lewontin intended these resonances between evolutionary theory and neoclassical economics in this paper, given his Marxist commitments.

themselves make an immediate individual contribution to fitness, but are rather only there because they are necessary to any structure in which other, more obviously useful, traits or elements occur together.

The question now is whether Kant's panfunctionalism commits him to the Panglossian paradigm, i.e. whether Kant indeed wishes to claim that every part of an organism makes an immediate individual contribution to the usefulness of the organism. I believe not, because on Kant's theory, a function is not necessarily something as straightforward as an individual contribution to the fitness of the organism. Indeed, what may trouble many readers of Kant is that his notion of function is not identifiable with the familiar concept of function known from evolutionary theory. Instead, Kant focuses on the intimate interconnection between the parts of an organism, and their mutual productive and formative capacities. Thus, a part of an organism can be functional if it makes a structural contribution to the organism as well, i.e. if it is a condition for the presence and proper functioning of other parts. Ironically, this is the very notion of functionality that Lewontin and Gould seem to endorse in their paper as well, since they suggest that we should judge organisms as wholes, with the historic themes of developmental morphology and *Bauplan* in mind, themes that are more closely connected with the Kantian tradition in the life sciences. In a way, then, Lewontin and Gould's account does not target the Kantian notion of function as of a piece with the notion that should be removed from evolutionary biology – it suggests that we *reinsert* the Kantian notion of function in evolutionary biology.

In stressing the centrality of the organism in Kant's philosophy of the life sciences, I may have already forced a negative answer to the question of whether ecologies can be considered as natural purposes. Yet, it is not because Kant focused on organisms, and obviously preferred organisms as paradigms, that he *could* not allow systems other than organism to be judged as natural purposes. After all, Kant himself notes that the admission of internal purposiveness allows for a certain slippage which lets us consider nature as a whole as a purposive system, in which everything is purposive and functional. Kant's reticence towards this idea is understandable, because it is far more difficult and misleading to assess the contributions of a thing to an ecology rather than to an organism, primarily because it is so difficult to *identify* an ecology or an ecological niche. Given the amount of philosophical work done by individuality in Aristotelian and Kantian approaches to the life sciences, this is a serious challenge to the recognition of ecological systems as natural purposes. But maybe recent developments in the theory of

the life sciences have removed some of the challenges by developing a more flexible account of biological individuality<sup>17</sup>.

## 4.3 Assuaging the Antinomy of Teleological Judgment

### 4.3.1 Presentation of the Antinomy

Like every Kantian Critique, the “Critique of the Teleological Power of Judgment” contains a “dialectic”, i.e. a section which deals with the errors to which our cognitive capacities are liable to lead because of their own nature, and which must be revealed as erroneous. In other works, however, the errors are due to reason, and its specific

---

<sup>17</sup> I am thinking specifically of two strands of thought on the individuation of organisms that have arisen in the past 25 years. The first is the conception of a developmental system proposed by Griffiths and Gray (1994), who note that the following thought is perhaps the most surprising result of their account: “Organism and environment are both evolving as an effect of the evolution of differentially self-replicating life cycles. Life cycles still have fitness values, but these are interpreted, not as a measure of correspondence between the organism and its environment but as measures of the self-replicating power of the system. Fitness is no longer a matter of “fittedness” to an independent environment.” (Griffiths & Gray 1994: 300-301). Thus, the biological systems that they individuate are neither genes nor organisms. The other interesting reconception of biological individuality is that suggested by Thomas Pradeu’s study of the philosophical relevance of immunology. Pradeu argues that immunology is concerned with investigating the problem of biological individuality. This is because “the immune system, with its surveillance activity, defines what is accepted or rejected by the organism. A criterion of immunogenicity thus constitutes a criterion of inclusion: the distinction between entities that are interconnected and form a whole as constituents of the organism and those that are rejected is carried out by the immune system. In other words, immunology allows for an understanding of the living thing’s spatial boundaries and by extension determines which entities constitute its components.” (Pradeu 2012: 240). Here, the remarkable result is that the organism is “heterogenous”, i.e. that “an organism is made of countless foreign entities; it is never constructed in a purely endogenous way. This heterogeneity can be illustrated by the functional role of certain symbiotic bacteria in their host. Each human being is made up of symbiotic bacteria numbering at least ten times those that are its “own” cells, in the sense of cells originating from the egg cell. The majority of these symbiotic bacteria live in the intestine. Most are obligatory symbionts, which means that they cannot survive outside the host, and that without them the host cannot survive either. They play indispensable physiological (functional) roles, particularly in digestion and immunity.” (Pradeu 2012: 247). Of these two perspectives, Pradeu’s is closer to Kant, because it conceives of biological individuality through physiological integration rather than through evolution and selection. Nevertheless, the account of Griffiths & Gray strongly suggests that the integration of an organism with its environment, with its ecology, might be stronger than suspected in much the same way that the connection between an organism and its symbionts is stronger than traditional conceptions of biological individuality seem to suppose.

internal structure, whereas here Kant speaks of a dialectic of the *power of judgment*. We will deal in detail with the capacity of the power of judgment to end up in such issues in the next chapter. In this section, we will focus on the antinomy of the teleological power of judgment and its resolution. In this subsection, I will discuss Kant's presentation of the antinomy in §70, and then offer my interpretation of Kant's announcement of the general strategy of the resolution of the antinomy in §71. In the next subsection, I will discuss Kant's presentation of the different theories of teleology in §72, his assessment of these theories in §73 and his general conclusion regarding the puzzle in §74. In the final subsection of this chapter, I will discuss Kant's resolution of the antinomy as it appears in §§75-78.

In §70 of the Dialectic of the "Critique of the Teleological Power of Judgment", Kant presents the antinomy as one that runs between two maxims:

The **first maxim** of the power of judgment is the **thesis**: All generation of material things and their forms must be judged as possible in accordance with merely mechanical laws.

The **second maxim** is the **antithesis**: 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). (AA V: 387)

There are two interpretative difficulties with this presentation of the antinomy, both of which are revealed by the immediately preceding passage, which is responsible for the standard interpretation of Kant's position here up until Peter McLaughlin's 1990 challenging reinvigoration of the study of the "Critique of the Teleological Power of Judgment". The passage runs:

Now in the case of this contingent unity of particular laws the power of judgment can set out from two maxims in its reflection, one of which is provided to it by the mere understanding *a priori*, the other of which, however, is suggested by particular experiences that bring reason into play in order to conduct the judging of corporeal nature and its laws in accordance with a special principle. It may then seem that these two sorts of maxims are not consistent with each other, thus that a dialectic will result that will make the power of judgment go astray in the principle of its reflection. (AA V: 386-387)

Kant has been read as saying here that, first, the antinomy is between a principle of the understanding *a priori*, and a principle of the reflective power of judgment, and second, that it is only seemingly there. Indeed, this has been the standard reading (for this, see McLaughlin's (1990: 137-142) invaluable analysis of different interpretations of the Antinomy). The first idea is then supported by the claim that the first maxim, that of mechanism, is simply the principle of causality as expressed in the Analytic of the *Critique of Pure Reason*, whereas the second maxim is a heuristic one adopted for the sake of understanding particular natural entities through reflective judgment. This delivers

the first element to Kant's supposed answer to the antinomy, namely that the two principles are not on the same footing. The second element is that these principles are merely regulative, and not constitutive, and therefore do not truly conflict. The latter reading is then supported by the fact that Kant immediately goes on to contrast these two maxims with the two principles as constitutive, and the contradiction into which they are supposed to run:

Now if one were to transform these regulative principles for research into constitutive principles of the possibility of the objects themselves, they would run:

**Thesis:** All generation of material things is possible in accordance with merely mechanical laws.

**Antithesis:** Some generation of such things is not possible in accordance with merely mechanical laws.

In this latter quality, as objective principles for the determining power of judgment, they would contradict one another, and hence one of the two propositions would necessarily be false; but that would then be an antinomy, though not of the power of judgment, but rather a conflict in the legislation of reason. However, reason 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. (AA V: 387)

But there are several serious objections to the identification of either supposed element of the resolution of the Antinomy. First of all, as McLaughlin has pointed out, readings depending on either element lead to the conclusion that Kant has resolved the antinomy, or at least believes to have done so, in §70. This raises the difficult interpretative issue of why Kant ended up writing 8 more paragraphs. Furthermore, if the solution of the antinomy consists in showing that there really is no antinomy as long as we bear in mind that we are dealing with reflective judgment rather than with determinate judgment or understanding, then it becomes mysterious why Kant elected to add an antinomy at all. McLaughlin has shown that, in order to explain Kant's expansiveness on a moot issue, many interpreters end up massively violating the principle of charity by stating that there really was no reason to write any of this, and Kant merely adds it for reasons of architectonic symmetry. This is of course highly problematic, not only because of the violation of the principle of charity, but also because it is simply wrong that Kant believes the distinctness of topics cannot overrule the symmetry of architectonics. After all, neither the *Critique of Practical Reason* nor the two parts of the *Critique of the Power of Judgment* contain an aesthetic. Moreover, Kant says that, because of the nature of the power of judgment, its critique can have no doctrinal part, and hence no doctrine of method. Thus, Kant is not above abandoning architectonic symmetry in textual and argumentative structure if it is demanded by his

theory on the specific status of different faculties. Thus, if it were the case that the teleological power of judgment cannot lead to an antinomy because of its merely regulative status, Kant could perfectly have said this.

Besides this more general objection, there are textual reasons for dismissing standard reasons. I will discuss three such reasons, namely that the principles cannot be read as stemming from different faculties, that the antinomy is not resolved by removing the blatant contradiction of the determinative and constitutive version, and that Kant clearly intends these paragraphs as a statement of the problem and a preparation for the solution rather than as a dismissal of the problem and *ipso facto* presentation of the solution.

First of all, it is not the case that Kant regards the regulative version of the antinomy as between a constitutive, determinative principle and a regulative one, for in that case, he would have to say that the first principle is provable by reason. After all, if the first maxim comes down to the principle of causality, then not only is it provable, Kant would have claimed to have proven it in the Analytic of the *Critique of Pure Reason*. But Kant insists that “reason can prove neither the one nor the other of these fundamental principles” (AA V: 387). Thus, whatever the content of the first maxim, it is not identical with the category of causality. Kant does say, of course, that the first maxim is “provided to it by the mere understanding *a priori*” (AA V: 386), but this does not have to imply that it is a principle of the mere understanding *a priori*. Given the straightforward contradiction (within the distance of a page) that would occur if we were to draw this implication, a different reading of how the understanding *a priori* would provide the first maxim would be desirable. In the next chapter, we will provide such a reading, but for now we will merely note the unsatisfactory nature of the opposite thesis.

Second, the removal of the contradiction between the two principles does not fully remove the antinomy. In the case of the opposition of the two principles read as constitutive, the two claims obviously contradict each other, since the first thesis says that all A are p, whereas the second thesis says that some A are not p. Kant suggests that by transforming the two theses from constitutive principles to regulative maxims, this contradiction is removed. I assume he believes this to be the case because the constitutive version makes generalizing statements about the objects, whereas the regulative version says something about the judging subject, and stating that a subject is under conflicting demands does not imply a contradiction, for not only are double binds logically possible, they are extremely common and pervasive in the concrete lives of human beings.

But Kant is here not speaking of the concrete lives of human subjects and the conflicting personal and societal demands and obligations under which they find themselves – he is talking about the subject of cognition and the transcendental constraints on his cognitive activity. The conflicting demands of reflective judgment are therefore not a matter of the vicissitudes of our psychological life, but rather an *a priori*

condition of finite cognition. The conflict is, like all antinomies, self-incurred. How is one then to proceed in judging according to conflicting maxims? This, I believe, is the problem sketched by Kant in §71, and which he attempts to resolve in the subsequent paragraphs of the “Antinomy of the teleological power of judgment”.

In §71, which Kant titled “Preparation for the resolution of the above antinomy”, Kant presents what I take to be the general strategy of his resolution:

Whether, therefore, the productive capacity of nature may not be as adequate for that which we judge as formed or combined in accordance with the idea of ends as well as for that which we believe to need merely the machinery of nature, and whether in fact things as genuine natural ends (as we must necessarily judge them) must be based in an entirely different kind of original causality, which cannot be contained at all in material nature or in its intelligible substratum, namely, an architectonic understanding: about this our reason, which is extremely limited with regard to the concept of causality if the latter is supposed to be specified a priori, can give us no information whatever. – However, with respect to our cognitive faculty, it is just as indubitably certain that the mere mechanism of nature is also incapable of providing an explanatory ground for the generation of organized beings. It is therefore an entirely correct fundamental principle for the reflecting power of judgment that for the evident connection of things in accordance with final causes we must conceive of a causality different from mechanism, namely that of an (intelligent) world-cause acting in accordance with ends, no matter how rash and indemonstrable that would be *for the determining power of judgment*. (AA V: 388-389)

As is to be expected, Kant takes as a premise for this argument the idea that we cannot actually *prove* that some objects cannot be explained according to the mere mechanism of nature, for this would make the principle of teleology constitutive. This means that we can never determinately identify an object as a natural purpose, even though there are objects which we must necessarily judge as natural purposes. The problem, however, is that natural ends “must be based in an entirely different kind of original causality, which cannot be contained at all in material nature or in its intelligible substratum, namely, an architectonic understanding”. With this statement, Kant repeats what I have argued to be the upshot of the Analytic, namely that natural purposes cannot be understood through either the merely mechanical causality of nature or the causality of intentional agency. What is now peculiar is that Kant nonetheless concludes that we *should* judge natural purposes according to the principle of intentional action. I will argue, in the next two subsections, that Kant’s resolution to the antinomy consists precisely in suggesting that we should heuristically understand natural purposes not according to the incomprehensible causality that is proper to them, but according to the non-mechanical but comprehensible causality of intentional agency, because the latter can be harmonized with mechanism in a way in which the former cannot. The

mechanicist paradigm, after all, allows for the compatibility and the neat distinctness of the mechanical and the intentional levels of explanation, and allows the latter to inform the former as a heuristical means. Kant's solution, then, is that we adopt the Boylean position on final causes heuristically in order to harmonize our otherwise massively conflicting commitments to mechanicism, which denies the productivity of nature, and the doctrine of natural teleology, which insists on the productivity of nature.

### 4.3.2 Candidates for Heuristics

In this subsection, I will argue that, in §§72-74, Kant develops an argument for his choice of intentional causality as the best and all only admissible candidate for a heuristic approach to living systems that does not massively conflict with the maxim of mechanicism. It is in the light of this task that we should read the following passage from §72:

Insofar as we would call the procedure (the causality) of nature a technique, on account of the similarity to ends that we find in its products, we would divide this into *intentional technique* (*technica intentionalis*) and *unintentional technique* (*technica naturalis*). The former would mean that the productive capacity of nature in accordance with final causes must be held to be a special kind of causality; the latter that it is at bottom entirely identical with the mechanism of nature, and that the contingent coincidence with our concepts of art and their rules, as a merely subjective condition for judging nature, is falsely interpreted as a special kind of natural generation. (AA V: 390-391)

This passage is bound to lead to misinterpretations, for here Kant seems to be saying that every causality that is not intentional is ultimately identical to mechanical causality, whereas I have been arguing that he sees natural purposiveness as a non-mechanical and non-intentional causality. As a result, it is tempting to read here that ultimately, Kant's notion of natural purpose is intended to allow for the fact organisms can be fully mechanical, although it is heuristically useful to regard them, at least for the time being<sup>18</sup>, as non-mechanical. I believe this is a mistake, however, as is revealed by the context of the passage. In this paragraph, Kant will list four historical alternative stances towards teleology in nature. The idea that the option between unintentional and intentional purposiveness collapses into the option between denying and admitting

---

<sup>18</sup> For example, Angela Breitenbach (2008) has argued that, for Kant, the idea of teleology is merely one that we must adopt until a proper mechanicism reduction is available. The idea behind this is that we need to have a preliminary concept to identify the reduced concept (that of natural purpose) to be able to reduce it to the reducing concept (that of mechanical system).



purposiveness is an underlying assumption of the traditional list, not of Kant's own approach. Indeed, he will stress that none of these approaches actually capture the idea of purposiveness that he seeks to accommodate. I therefore suggest we read Kant not as subscribing to that dichotomy, but rather as showing which options are open once one accepts it.

Before discussing the four positions, I would first like to offer a table which shows how Kant clarifies the various positions:

<b>Purposiveness is</b>	<i>Unintentional</i>	<i>Intentional</i>
<i>Due to the physical</i>	Epicureanism	Hylozoism
<i>Due to the hyperphysical</i>	Spinozism	Theism

Kant presents the two positions in the first column as follows:

The **idealism** of purposiveness (I always mean objective purposiveness here) is now either that of the **accidentality** or of the **fatality** of the determination of nature in the purposive form of its products. The first principle concerns the relation of matter to the physical ground of its form, namely the laws of motion; the second concerns the **hyperphysical** ground of matter and the whole of nature. The system of **accidentality**, which is ascribed to Epicurus or Democritus, is, if taken literally, so obviously absurd that it need not detain us; by contrast, the system of fatality (of which Spinoza is made the author, although it is to all appearance much older), which appeals to something supersensible, to which our insight therefore does not reach, is not so easy to refute, since its concept of the original being is not intelligible at all. But this much is clear: that on this system the connection of ends in the world must be assumed to be unintentional (because it is derived from an original being, but not from its understanding, hence not from any intention on its part, but from the necessity of its nature and the unity of the world flowing from that), hence the fatalism of purposiveness is at the same time an idealism of it. (AA V: 391-392)

The idealism of purposiveness takes that which creates the illusion of purposiveness in nature to be fully explainable without purposiveness, since there is no, nor can there be, intentional agency involved. Kant distinguishes between two such positions, namely one that explains the seeming purposiveness in terms of blind chance, and one that explains it in terms of necessity. A further complication is that Kant believes this to coincide with a different opposition, one between physical and hyperphysical grounds, which is far less obvious. I will first discuss the distinction between blind chance and fatality and then offer a way of identifying with the distinction between the physical and the hyperphysical.

Kant presents the first kind of idealism of purposiveness as "Epicurus's kind of explanation, on which the difference between a technique of nature and mere

mechanism is completely denied, and blind chance is assumed to be the explanation not only of the correspondence of generated products with our concepts of ends, hence of technique, but even of the determination of the causes of this generation in accordance with laws of motion, hence of their mechanism” (AA V: 393). Although he does not say so explicitly, it becomes clear from the way in which he contrasts this position with that of fatality that epicureanism intends to explain the origin of the apparent contingency of natural purposes with regard to general laws. This contingency, the epicurean would say, is not due to the involvement of an intentional agency in their production, but rather to the element of randomness that blind chance inserts. The apparent purposiveness is thus indeed contingent, but blindly and not purposively contingent.

In contrast, the fatalism of purposiveness does not allow for such an element of blind chance, since it does not allow for any chance, coincidence or indeterminacy:

Spinoza would suspend all inquiry into the ground of the possibility of the ends of nature and deprive this idea of all reality by allowing them to count not as products of an original being but as accidents inhering in it, and to this being, as the substratum of those natural things, he ascribes not causality with regard to them but merely subsistence, and (on account of the unconditional necessity of this being, together with all natural things as accidents inhering in it), he secures for the natural forms the unity of the ground that is, to be sure, requisite for all purposiveness, but at the same time he removes their contingency, without which no *unity of purpose* can be thought, and with that removes everything *intentional*, just as he removes all understanding from the original ground of natural things.  
(AA V: 393)

The approach identified as Spinoza’s thus does not explain the illusion of purposiveness by showing how contingency can occur in nature, but rather by showing how the world and the phenomenon can be understood as from a unitary ground, not an aggregate of parts and processes. But the spinozist does not explain the unity of purpose, for he focuses not on the illusion of purposiveness, but rather on the unity of the organism. The problem of course is that the organism does not have this unity because of its own nature, but rather because its subsistence is an expression of the power of the one substance in which it inheres.

This brings us to the distinction between physical and hyperphysical grounds of the alleged purposiveness in nature. For epicureanism, nature has general principles, and organisms are entirely the product of these general principles and the blind chance that operates within them. Organisms are therefore regarded as being due to natural principles. On the spinozist position, on the other hand, the alleged purposiveness is not due to the inner workings of nature, but due to the ground of nature as a whole.

In order to understand the importance of the distinction between the two positions, however, we will have to see how the distinction works out in the case of the realism of purposiveness, for only there does Kant fully articulate it. One reason why Kant may not

fully develop this distinction in the case at hand is because of the unsatisfactory nature of both accounts for his purposes. They are unsatisfactory not because they deny purposiveness, but because they cannot account for the illusion of purposiveness they posit: that is, they cannot show how, in spite of the fact that there is no purposiveness at work in nature, the concept of purposiveness is so strongly suggested by some natural occurrences. In the case of epicureanism, this is because blind chance may account for the contingency of a thing with respect to the general principles, of nature, but not for the unity this thing exhibits in its contingency. If it were merely due to blind chance, we would expect this chance to be equally at work within the organism, but the latter's functional organization belies such randomness. Spinozism, on the other hand, explains not the contingency with respect to general principles, but rather the unity exhibited by modes (the striving making up, or made up by, their individual essences)<sup>19</sup>. Such unity, however, is not functional unity, and thus is not the kind of purposiveness that we would be tempted to postulate.

The importance of this analysis is that it shows us not why epicureanism and spinozism are wrong, but that they turn out to be bad candidates for the harmonization of mechanism and teleology. Both approaches can initially be regarded as such harmonization, because they are supposed to explain why the appearance of purposiveness in nature can be squared with the absence of purposiveness. Kant concludes that neither of these idealisms of purposiveness can explain why nature would falsely appear to be purposive, and therefore cannot harbor the teleological principle.

Kant next considers the two forms of the realism of purposiveness:

The **realism** of the purposiveness of nature is also either physical or hyperphysical. The **first** bases ends in nature on the analogue of a faculty acting in accordance with an intention, the **life of matter** (in it, or also through an animating inner principle, a world-soul); and is called **hylozoism**. The **second**

---

<sup>19</sup> There have been several authors who suggested that Spinoza's concept of conatus ultimately committed him to some form of teleology. Nietzsche, for instance, wrote, in *Beyond Good and Evil* (2002: 15): "In short, here as elsewhere, watch out for *superfluous* teleological principles! – such as the drive for preservation (which we owe to Spinoza's inconsistency)". Don Garrett (1999) has suggested that we resolve the inconsistency by reading Spinoza as allowing for unintentional teleology in the form of the conatus, and rejecting only intentional teleology. This would make him more akin to Aristotle and Kant. Whether this reading is desirable or not, the general reception of Spinoza warrants Kant's interpretation of Spinozism as anti-teleological. Moreover, Kant would seem to have discounted the teleological properties of the conatus because he regarded the conatus as nothing more than a modal expression of God's own non-teleological power of subsistence as *causa sui* (as is clear from his comments on the issue). This reading may be incorrect, but it is neither crude nor obviously unacceptable.

derives them from the original ground of the world-whole, as an intentionally productive (originally living) intelligent being; and it is **theism**. (AA V: 392)

These two positions are not distinguished by the aspect of the appearance of purposiveness that they explain, or neglect, for they both fully allow for both aspects of purposiveness. They can, however, be distinguished by the distinction between the physical and the hyperphysical ground of purposiveness, to the extent that one cannot help but feel that this distinction is mostly applicable to the realism of purposiveness, and only derivatively to the idealism of purposiveness. Hylozoism explains purposiveness in nature by ascribing to matter itself or to nature itself an animating capacity for organization, whereas theism explains it by positing, apart from nature, a transcendent, external intelligent agent responsible for organizing nature. Thus, like epicureanism, hylozoism locates the ground of purposiveness in nature itself, although in this case this ground is intentional rather than blind. Theism, on the other hand, like Spinozism, locates that capacity in the hyperphysical ground of nature, although in this case the hyperphysical ground is transcendent rather than immanent and intentional rather than necessitarian<sup>20</sup>.

After having described and categorized the possible positions, Kant goes on to assess them, beginning with hylozoism, which he resists vehemently:

the possibility of a living matter (the concept of which contains a contradiction, because lifelessness, *inertia*, constitutes its essential characteristic), cannot even be conceived; the possibility of an animated matter and of the whole of nature as an animal can be used at all only insofar as it is revealed to us (for the sake of an hypothesis of purposiveness in nature at large), in experience, in the organization of nature in the small, but its possibility can by no means be understood a priori. There must therefore be a circle in the explanation if one would derive the purposiveness of nature in organized beings from the life of matter and in turn is not acquainted with this life otherwise than in organized beings, and thus cannot form any concept of its possibility without experience of them. Hylozism thus does not accomplish what it promises. (AA V: 394-395)

Kant's rejection of hylozoism is easily misinterpreted as a definitional issue, for he claims that it is a contradiction. Note, however, that the argument is more complex than

---

<sup>20</sup> I say necessitarian rather than necessitated because, strictly speaking, Spinoza's God is free and not necessitated, since he is not determined by anything external to him. Nevertheless, God is not intentional because he necessarily acts in the way he does precisely because he is fully determined by his own nature. It may have been the case that this nuance was lost on Kant, much like it was lost on many of those who read or heard of Spinoza's theories. However, it is important to see that Kant's argument does not rely on Spinoza's God being necessitated, but rather on his being unintentional.

that. The first argument basically states that one cannot ascribe life to matter a priori, whereas we could ascribe inertia to it. But he does not therefore exclude that we might for empirical reasons decide to ascribe further properties to matter. I suggest we read this passage as an allusion to Newtonianism, for indeed Newton specified matter in terms of inertia in the principles, and nevertheless went on to allow for gravity. Of course, Kant would prefer to think of gravity as a force and not a property of matter, but even for those who would think of it in the latter way, the parallel between vitality and gravity would not hold. As Kant notes, there is a circularity involved in deciding that the organization of living beings is due to the life of matter, and seeing the life of matter as manifesting itself only in the organization of living beings. I take it he means to say that it is *ad hoc*. After all, gravity is a general property of matter (on the relevant interpretation) that exhibits itself in every physical system and interaction, whereas vitality would only manifest itself in living systems. For such an *ad hoc* posit, the Newtonians' argument that gravity is not mysterious because it is simple and ubiquitous would not bring solace.

This criticism against hylozoism is a familiar one in Kant's oeuvre, for, as we saw in section 2.3.1, he already raised it in 1763 when criticizing the Newtonian analogy as used by Buffon and Maupertuis in their theories of generation. This reveals an important aspect of the argument, namely that it was directed against those who would expand the list of general principles allowed in mechanicism in order to accommodate purposiveness. Such attempts at expansion of mechanicism, Kant believed, were doomed to fail because the kind of principles that could be general enough to be acceptable in mechanicism could not explain the specificity of purposiveness, and the kind of principles that could explain the specificity of purposiveness would be too specific to be acceptable in mechanicism. Here, too, we see Kant assessing hylozoism not in itself, but as a possible way of harmonizing teleology and mechanicism.

The final option Kant considers is theism, namely the familiar theory that (some) natural entities exhibit purposiveness because they are intentionally designed. Like all other positions, it cannot provide a ground for "dogmatically establishing the possibility of natural ends as a key to teleology" (AA V: 395), but Kant believes that it is superior to all others because it is not as obviously unacceptable as hylozoism. For this reason, he concludes the paragraph as follows:

for us there remains no other way of judging the generation of its products as natural ends than through a supreme understanding as the cause of the world. But that is only a ground for the reflecting, not for the determining power of judgment, and absolutely cannot justify any objective assertion. (AA V: 395)

In light of Kant's constant criticism of the idea that natural purposiveness is best understood as intentional, this passage would be bizarre, because here he seems to conclude that design is the best conception after all. My answer to this puzzle will

probably have become clear by now: Kant does not embrace design because it is the best theory of natural purposiveness, but because it is the best candidate for a heuristic to inquire into natural purposes within the constraints of mechanism. After all, Kant's intention here is to find a way out of our necessity of regarding organisms in terms of the problematic concept of a natural purpose, which

[a]s a concept of a **natural product** [...] includes natural necessity and yet at the same time a contingency of the form of the object (in relation to mere laws of nature) in one and the same thing as an end; consequently, if there is not to be a contradiction here, it must contain a basis for the possibility of this thing in nature and yet at the same time a basis of the possibility of this nature itself and its relation to something that is not empirically cognizable nature (supersensible) and thus is not cognizable at all for us, in order to be judged in accordance with another kind of causality than that of the mechanism of nature, if its possibility is to be determined. (AA V: 396)

The direct conflict of a natural purpose as both natural and purposive is assuaged under the assumption of design because, although the natural purpose is regarded entirely as a natural entity obeying the mechanical principles, nature itself, specifically its organization and its structure, is regarded as intentionally designed. This allows for a resolution of the antinomy by neatly partitioning the workings and the structure of nature between the two principles: the workings of an organism are to be understood entirely according to mechanical principles, but the structure which constrains these mechanical principles is not itself understood as a product of the working of nature, but as due to the (original) structuring of nature by a creator. But this way of regarding natural purposiveness is necessary only "because of the peculiar constitution of my faculties" (AA V: 397).

### 4.3.3 The Resolution of the Antinomy

In this subsection, I will present §§75-78 of the *Critique of Judgment* as the actual solution to the antinomy of the teleological power of judgment. As we will see, §75 explains why the theistic conception of teleology can be adopted without thereby altering the merely regulative (and not constitutive) status of the two maxims of the teleological power of judgment. §76 will not be discussed in detail, because it is basically a digression, or rather a reminder of the meaning of the concept of a regulative principle or idea. §77, on the other hand, is essential because it tells us which limitations of human cognition make this solution necessary and why. §78, finally, shows how the suggested position solves the antinomy.

Kant opens §75 with the invaluable reminder that the adoption of a theistic conception of teleology is not to be taken as constitutive:

To say that the generation of certain things in nature or even of nature as a whole is possible only through a cause that is determined to act in accordance with intentions is quite different from saying that *because of the peculiar constitution of my cognitive faculties* I cannot judge about the possibility of those things and their generation except by thinking of a cause for these that acts in accordance with intentions, and thus by thinking of a being that is productive in accordance with the analogy with the causality of an understanding. In the first case I would determine something about the object, and I am obliged to demonstrate the objective reality of a concept that has been assumed; in the second case, reason merely determines the use of my cognitive faculties in accordance with their special character and with the essential conditions as well as the limits of their domain. The first principle is thus an objective fundamental principle for the determining, the second a subjective fundamental principle merely for the reflecting power of judgment, hence a maxim that reason prescribes to it. (AA V: 397-398)

My discussion in the previous subsection allows us to see why Kant insists so much on this distinction and why he can. The theistic position is not itself suggested by the necessity of judging some natural entities as natural purposes, since natural purposiveness itself is not adequately covered by intentional or practical causality. The problem is rather that no concept of causality adequate to capture natural purposiveness is comprehensible for our form of cognition. This is why we adopt the theistic stance, namely because it has the merit of allowing us to capture some important features of purposiveness whilst still operating within the mechanistic framework forced upon us by the limitations of our knowledge. As a result, we can never conclude from the need to admit purposiveness in nature to the need to admit the intentional causation of a supersensible being. The latter principle is necessary only for our empirical scientific investigations into organized beings (AA V: 398). Once we have adopted this principle, however, it becomes natural to regard teleology and theology as necessary for one another:

Now the concept of a thing whose existence or form we represent as possible under the condition of an end is inseparable from the concept of its contingency (according to natural laws). Hence natural things which we find possible only as ends constitute the best proof of the contingency of the world-whole, and are the only basis for proof valid for both common understanding as well as for philosophers of the dependence of these things on and their origin in a being that exists outside of the world and is (on account of that purposive form) intelligent; thus teleology cannot find a complete answer for its inquiries except in a theology. (AA V: 398-399)

It is perhaps too tempting to once again read Kant as agreeing unambiguously with this pervasive idea, namely that the full understanding of teleology would require, and in its

turn imply, theology. The next passage immediately reveals, however, that Kant is aware of the lack of fit or convergence:

But what does even the most complete teleology prove in the end? Does it prove anything like that such an intelligent being exists? No; it proves nothing more than that because of the constitution of our cognitive faculties, and thus in the combination of experience with the supreme principles of reason, we cannot form any concept at all of the possibility of such a world except by conceiving of such an intentionally acting supreme cause. (AA V: 399)

Here Kant is saying that it we cannot conclude from the contingency of the world-whole that it is intelligently designed, since it is only because of the peculiar constitution of our cognitive faculties that we need to conceive of this contingency as due to intentionality. Mind that Kant is *not* saying that it is only due to the peculiar constitution of our cognitive faculties that we need to regard the world as containing contingency. The problem is not whether some aspects in the world need to be understood as mechanically inexplicable or not, but rather over whether this mechanical inexplicability is due to intelligent design or not. Thus, whereas we are somehow required to admit to the occurrence of ends in nature,

we do not actually observe ends in nature as intentional, but merely add this concept as a guideline for the power of judgment in reflection on the products of nature, they are not given to us through the object. It is even impossible for us to justify a priori the assumption of the objective reality of such a concept. There is thus left nothing but a proposition resting only on subjective conditions, namely those of a reflecting power of judgment appropriate to our cognitive faculties, which, if one were to express it as objectively and dogmatically valid, would say: There is a God; but all that is allowed to us human is the restricted formula: We cannot conceive of the purposiveness which must be made the basis even of our cognition of the internal possibility of many things in nature and make it comprehensible except by representing them and the world in general as a product of an intelligent cause (a God). (AA V: 399-400)

The main benefit of this reading is that it can help us make sense of a famous passage from Kant which would otherwise be deeply problematic:

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; 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 underlying their generation would be presumptuous: for how could we know that? Probabilities count for nothing here, where judgments of pure reason are at stake. – Thus we cannot make any objective judgment at all, whether affirmative or negative, about the proposition that there is an intentionally acting being as a world-cause (hence as an author) at the basis of what we rightly call natural ends. (AA V: 400)

To see why this passage is incomprehensible on traditional readings, observe that Kant is here saying both that it is impossible that nature can be understood fully mechanically, and that it is impossible to prove that this feature of nature cannot be due to a causality that is non-intentional. If Kant equated non-mechanical causation with intentional causation, this point would be contradictory, for he would be saying that it is impossible to fully understand nature merely mechanically, and that it is impossible to prove that nature cannot be fully understood merely mechanically. Thus, we cannot, on pain of contradiction or of ascribing to Kant a blatant contradiction, embrace this interpretation. And yet this interpretation is ubiquitous.

Peter Mclaughlin's interpretation of the resolution of the antinomy clearly endorses the interpretation leading to the contradiction, since it runs as follows:

The resolution of the conflict between the general necessity and the occasional impossibility of mechanical explanation is thus the following: We must judge all natural things mechanistically because for us only mechanical objects can be explained. If we are unable to conceive of a particular object of experience as naturally mechanical, we must judge it as an artificial mechanism that was intended by some understanding. *This is not because such an understanding exists, nor because the thing is not really mechanical (if we knew all empirical laws we might be able to conceive it as mechanical)*, but because we cannot otherwise conceive the apparent causal dependency of the parts on the whole. (Mclaughlin 1990: 168; my stress)

In the passage above, I have stressed the sentence to which I unambiguously object, for it ascribes to Kant the idea that we cannot determinately state that these things are mechanically inexplicable. First of all, this reading ends up ascribing to Kant the contradiction I just offered. Secondly, Kant does state that nature cannot be understood merely mechanically, since he insists that the concept of a natural purpose implies mechanical inexplicability, and that in teleological judgment we rightly call some things natural ends. What he believes is a bridge too far is the admission that this mechanical inexplicability is due to intentional causation. This passage, however, does not speak of mechanical causation, but rather of unintentional causation. We can thus avoid the contradiction by reading Kant as saying that we somehow know that nature is not fully mechanically explicable, but that we cannot know whether this is due to its being

designed, since it is always possible that this teleology is due to a non-intentional, natural teleology, precisely the kind of teleology he ascribes to natural purposes. The conclusion of the passage I have been discussing clearly supports such a reading, for Kant writes that “[t]hus we cannot make any objective judgment at all, whether affirmative or negative, about the proposition that there is an intentionally acting being as a world-cause (hence as an author) at the basis of what we rightly call natural ends.” In this sentence, Kant is stating that the issue is of whether we should ascribe an intentional ground to what we rightly call natural ends. But rightly calling something a natural end *is* rightly calling it mechanically inexplicable. The issue is thus not whether natural ends are in fact mechanically inexplicable, but rather whether the mechanically inexplicable natural ends are in fact intentionally designed.

Another reading that fails to account for this passage is that of James Kreines. Kreines (2005: 288) has argued that, for Kant, it is simply so that natural purposes themselves suggest intentional, intelligent design, the real question being whether something is rightly judged to be a natural purpose. The passage I just quoted cannot be harmonized with this reading, for here Kant is clearly saying that the decision that something is rightly called a natural purpose does not decide whether it is rightly called intentional. I feel that Kreines’ reading is suggested by the fact that somehow, Kant does not believe we can determinately judge things to be natural purposes either. This is an important issue in itself, and I will deal with it in the next chapter, where I will argue that it is the demand of specificity internal to the systematic drive of nature that compels us to regard organisms as natural purposes. My argument, however, will turn on a novel reading of the idea of systematicity. As a result, I will also be able to argue that this novel reading is interesting because it explains the peculiar status of the judgment of something as a natural purpose.

For now, Kant has been merely asserting, or suggesting, that the reason for adopting the theistic stance on natural purposes is the peculiar limitations of our cognitive capacities. Only in §77 does he clarify what he means by this. He introduces the issue with the following peculiar passage, which has been fateful in the reception of Kant’s argument:

What is at issue is therefore a special character of **our** (human) understanding with regard to the power of judgment in its reflection upon things in nature. But if that is the case, then it must be based on the idea of a possible understanding other than the human one (as in the *Critique of Pure Reason* we had to have in mind another possible intuition if we were to hold our own to be a special kind, namely one that is valid of objects merely as appearances), so that one could say that certain products of nature, as far as their possibility is concerned, **must**, given the particular constitution of our understanding, **be considered by us** as intentional and generated as ends, yet without thereby demanding that there actually is a particular cause that has the representation of an end as its determining ground,

and thus without denying that another (higher) understanding than the human one might be able to find the ground of the possibility of such products of nature even in the mechanism of nature, i.e., in a causal connection for which an understanding does not have to be exclusively assumed as a cause. (AA V: 405-406)

The move Kant makes here recurs throughout his work, namely that of contrasting our own cognition, which is irrevocably finite, with an infinite one, an “intuitive understanding”. The big initial puzzle is that Kant seems to be saying that, for such an alternative understanding, the purposiveness of nature might in fact be grounded in the mechanism of nature. This seems to run counter to his constant insistence that natural purposes are to be explained non-mechanically. We thus need to make sense of the kind of entity that would be able to explain the non-mechanical mechanically.

The first thing that we need to note is that the concept of an intuitive understanding always figures in Kant as a contrastive notion. It is tempting to read him as saying that there is, or can be, such an alternative understanding. One problem with this reading is that the contrastive idea is such that we can only conceive of it negatively, namely as something which would not be limited in the way we are. Positive characterizations of this form of knowledge are inherently deceitful. I believe part of the difficulty here is Kant’s suggestion that there the idea of intelligent design relies on a slippage from a negative to a positive characterization of the intuitive understanding.

The intuitive understanding is here first negatively characterized as that which does not have the contingency typical of the constitution of our understanding:

This contingency is quite naturally found in the **particular**, which the power of judgment is to subsume under the **universal** of the concepts of the understanding; for through the universal of **our** (human) understanding the particular is not determined, and it is contingent in how many different ways distinct things that nevertheless coincide in a common characteristic can be presented to our perception. Our understanding is a faculty of concepts, i.e., a discursive understanding, for which it must of course be contingent what and how different might be the particular that can be given to it in nature and brought under its concepts. 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 – a job that our understanding can accomplish only through the correspondence of natural characteristics with our faculty of concepts, which is quite contingent, but which an intuitive understanding would not need. (AA V: 406)

The contingency is due to our understanding being discursive, which means that we only ever grasp particulars and individuals through concepts. Such concepts never fully grasp, or can never be determinately admitted to fully grasp the particularities of that which is subsumed under it. This is because, if we want to think a particular or an individual, we must have a concept of it, since we cannot directly reason with it as presented in intuition. But it always remains a question of whether this concept is adequate to that which is subsumed under it. We usually grasp the particularities of particulars and individuals by developing more specific concepts under which to subsume them. As a result, there is always a certain gap between the concept and that thought under it, and the universal and the particular.<sup>21</sup>

Kant then determines the intuitive understanding as one for whom the grasping of particulars and individuals is not possible only through the more general concepts and principles under which they can be subsumed. For such an understanding, the harmony of the world and the conceptual system grasping it would not be contingent. This brings us one step closer to understanding Kant's peculiar point, since we saw that the problem of mechanism versus teleology was that the latter exhibits a specificity which runs counter to the generality of mechanical principles. Thus, this problem is that of a particular which is recalcitrant to subsumption under a universal, or a more general principle. For a discursive understanding, the threat of such recalcitrance is always at least in the background:

Our understanding thus has this peculiarity for the power of judgment, that in cognition by means of it the particular is not determined by the universal, and the latter therefore cannot be derived from the former alone; but nevertheless this particular in the manifold of nature should agree with the universal (through concepts and laws), which agreement under such circumstances must be quite contingent and without a determinate principle for the power of judgment. (AA V: 406-407)

To console us for this threat, however, we tend to conceive of an understanding for which this threat is not only non-existent, but which would also have the view of nature as the intelligible unity that we cognitively hope it to be:

Nevertheless, in order for us to be able at least to conceive of the possibility of such an agreement of the things of nature with the power of judgment (which we represent as contingent, hence as possible only through an end aimed at it), we must at the same time conceive of another understanding, in relation to which, and indeed prior to any end attributed to it, we can represent that agreement of

---

<sup>21</sup> This will become clearer in the next chapter, where I will deal with the maxims of homogeneity and specification.

natural laws with our power of judgment, which for our understanding is conceivable only through ends as the means of connection, as **necessary**. (AA V: 407)

We can assert that nature is such that it is amenable to being grasped by a discursive understanding. i.e. such that the particularities occurring in them can be understood on the basis of general or universal concepts or principles, only if we assume that nature was somehow meant to exhibit such comprehensibility. Moreover, we can only conceive of nature in this way if we conceive of an understanding for whom this harmony is immediate and unproblematic, and thus does not suggest that nature was meant to be cognized.

I would like to suggest that Kant offers this notion of an intuitive understanding as one for whom there would be no antinomy because it could conceive of a natural purpose as it is described in the “Analytic of the Teleological Power of Judgment”, whereas for us this concept is incomprehensible. An intuitive understanding could thus harmonize mechanism and teleology not because it sees that nature is inherently mechanical, but because it can reason otherwise than mechanically. Mechanical explanation, Kant insists throughout the “Critique of the Teleological Power of Judgment”, is the explanation of phenomena through the most general motive forces at work in nature alone. As a result, specificity and particularity are anathema to mechanicism, for they violate its principle of intelligibility. Mechanicism, it seems, is necessary for a discursive understanding because discursive understanding can only understand the particular through the general, and not through its own particularity. For an understanding that is not bound by such limitations, specificity can be directly, “intuitively” understood, and needn’t be reduced to more general principles.

This is in fact what we can read in the continuation of Kant’s argument:

Our understanding, namely, has the property that in its cognition, e.g., of the cause of a product, it must go from the **analytical universal** (of concepts) to the particular (of the given empirical intuition), in which it determines nothing with regard to the manifoldness of the latter, but must expect this determination for the power of judgment from the subsumption of the empirical intuition (when the object is a product of nature) under the concept. Now, however, we can also conceive of an understanding which, since it is not discursive like ours but is intuitive, goes from the **synthetically universal** (of the intuition of a whole as such) to the particular, i.e., from the whole to the parts, in which, therefore, and in whose representation of the whole, there is no **contingency** in the combination of the parts, in order to make possible a determinate form of the whole, which is needed by our understanding, which must progress from the parts, as universally conceived grounds, to the different possible forms, as consequences, that can be subsumed under it. (AA V: 407)

Kant thus characterizes the two forms of understanding, discursive and intuitive understanding respectively, as one which can only grasp the particular through general principles, and is always threatened by the possibility of the particular's recalcitrance to this subsumption under general principles, and one which can have a complete cognition of the particular in its particularity. For the latter, there is of course no contingency, for the particular or individual is understood through its own concept, and not through another concept. It is important to note that the problem here would not be of asking *how* the parts come to constitute this particular, for they are already understood as constituting this particular. Similarly, the question of whether the whole precedes the parts is a moot one, for the cause of the product need no longer be regarded as prior to it in the manner that an antecedent is prior to its consequent, since in an intuitive understanding the distinction between such moments of judgment is absent. Matters are different for discursive understanding, of course:

In accordance with the constitution of our understanding, by contrast, a real whole of nature is to be regarded only as the effect of the concurrent moving forces of the parts. Thus if we would not represent the possibility of the whole as depending upon the parts, as is appropriate for our discursive understanding, but would rather, after the model of the intuitive (archetypical) understanding, represent the possibility of the parts (as far as both their constitution and their combination is concerned) as depending upon the whole, then, given the very same special characteristic of our understanding, this cannot come about by the whole being the ground of the possibility of the connection of the parts (which would be a contradiction in the discursive kind of cognition), but only by the **representation** of a whole containing the ground of the possibility of its form and of the connection of parts that belongs to that. But now since the whole would in that case be an effect (**product**) the **representation** of which would be regarded as the **cause** of its possibility, but the product of a cause whose determining ground is merely the representation of its effect is called an end, it follows that it is merely a consequence of the particular constitution of our understanding that we represent products of nature as possible only in accordance with another kind of causality than that of the natural laws of matter, namely only in accordance with that of ends and final causes, and that this principle does not pertain to the possibility of such things themselves (even considered as phenomena) in accordance with this sort of generation, but pertains only to the judging of them that is possible for our understanding. (AA V: 407-408)

For a discursive understanding, the whole is to be understood according to the general principles of motion governing its parts. Such general principles and forces are the only ones which can be admitted as present prior to the constitution of the whole itself. In trying to understand the specification of the parts, however, as far as their relation and existence is concerned, we assume that the concept of the whole itself is also somehow

causally involved. The only way for us to do so is to assume that the concept of the whole is prior to the whole itself. Of course, it can be prior only as a representation in an intentional agent. An intuitive understanding, however, need not supplement its understanding with the concept of the particular, for it is already entertaining this concept; nor need it regard it as a pre-existing representation. It is therefore a consequence of the fact that our understanding is discursive, i.e. finite, that we understand the specificity at work in nature as external to nature, as due to an intentional agency restricting the otherwise too general working of nature.

The possibility of a different kind of understanding has two important functions at this stage of Kant's argument. First of all, it is meant to show that the two kinds of causality are only strictly opposed for our limited understanding, since for us to be a natural entity means to be explainable in terms of the general principle of nature, and to be purposive is to be explainable only by a non-natural causality. If we imagine a different kind of understanding that can have knowledge of natural entities otherwise than through general principles alone, we see that such an understanding could understand natural purposes without taking recourse to non-natural causality. Kant admittedly obscures this point at times by insisting that this intuitive understanding would understand natural purposes mechanically. But by this he always means that it would understand natural purposes solely through the natural activities of its parts, and not through the idea representation of a whole. Here, an equivocation takes place, because the mechanicism through which the intuitive understanding would understand natural purposes would not be what we mean by mechanicism, since it would not understand only through general principles. This terminological peculiarity is, I believe, due to a tendency, even in Kant, to regard mechanicism as the problem of eliminating supernatural explanations from natural philosophy. What Kant envisions is a mechanicism that would have eliminated all such explanations, but he admits that such a mechanicism would differ *toto caelo* from the one in which we are currently engaged.

The second important function of the idea of an intuitive understanding is to allow for the solution to the antinomy by harmonizing mechanicism and teleology. If another kind of understanding than ours is conceivable, then we cannot equate the way we need to think of the world with the way the world needs to be. Were we to in fact equate the two,

then the unity that constitutes the ground of the possibility of natural formations would be merely the unity of space, which is however no real ground of generating but only their formal condition; although it has some similarity to the real ground that we seek in that in it no part can be determined except in relation to the whole (the representation of which is thus the basis of the possibility of the parts). (AA V: 409)

The spatial, extensive configuration, is indeed the only admissible element in mechanical explanation besides general principles of motion. Thus, if the world were in fact spatially determined in the way we regard it, we would have no choice but to either dismiss teleology or allow it as another kind of causality working in the world. Luckily, Kant believes that the findings of transcendental idealism allow us to avoid this possibility. Through the adoption of the idea of a supersensible substrate,

since it is still at least possible to consider the material world as a mere appearance, and to conceive of something as a thing in itself (which is not an appearance) as substratum, and to correlate with this a corresponding intellectual intuition (even if it is not ours), there would then be a supersensible real ground for nature, although it is unknowable for us, to which we ourselves belong, and in which that which is necessary in it as object of the senses can be considered in accordance with mechanical laws, while the agreement and unity of the particular laws and corresponding forms, which in regard to the mechanical laws we must judge as contingent, can at the same time be considered in it, as object of reason (indeed the whole of nature as a system) in accordance with teleological laws, and the material world would thus be judged in accordance with two kinds of principles, without the mechanical mode of explanation being excluded by the teleological mode, as if they contradicted each other. (AA V: 409)

Kant here makes the bold suggestion to regard the empirical world as we know it to be merely the empirically knowable one, and to be subject to the limitations of our understanding, whilst equally holding it to be grounded in a substratum of which we cannot even conceive, and in which natural purposiveness can occur without intentional causation being involved. This supersensible substratum is characterized by Kant as the one to which we also belong, presumably in so far as we are moral agents. At first sight, it could seem that organisms can therefore be assumed to exhibit intentional free agency in the way that we assume ourselves to exhibit it. But it is clear from Kant's line of thought that this would be mistaken, since the whole point of a natural purpose is that it can exhibit purposiveness without intentionality. A better way to think of it may be to read the supersensible substrate as that which is governed by normativity, instead of – or possibly as well as – by non-normative causality. As we saw in section 4.2.3., it is indeed plausible to read Kant as thinking of natural purposes as unintentionally normative. This could indeed explain why Kant thinks grounding natural purposes in the supersensible allows us to avoid the apparent contradiction of natural normativity. We would in such a manner conceive of the world as if the kind of causality that he described in the *Analytic* were really at work in it, although such a world is incomprehensible to us. The result is that we proceed in our inquiries with the conviction that, although the antinomy between mechanical and teleological explanation is very much real for our form of cognition, it needn't be real at all, since



alternate causalities, alternate forms, are strictly speaking conceivable (though not comprehensible to us).

But conceiving of such a possibility is what allows us to take a regulative stance on the maxims; it is not what allows us to pursue them both simultaneously. Explaining how the two maxims can meaningfully be jointly pursued is the task of the final paragraph of the “Dialectic of the Teleological Power of Judgment”. There, Kant keeps insisting that, although the idea of a supersensible ground allows us to escape the contradiction between the two kinds of explanation, it does not show us how we can unite the two kinds of explanation in our scientific enterprises, since this supersensible ground is unknowable and can thus not be invoked in explanation:

Now of course the principle of the mechanism of nature and that of its causality according to ends in one and the same product of nature must cohere in a single higher principle and flow from it in common, because otherwise they could not subsist alongside one another in the consideration of nature. But if this objectively common principle, which also justifies the commonality of the maxims of natural research that depend upon it, is such that it can be indicated but can never be determinately cognized and distinctly provided for use in actual cases, then from such a principle there can be drawn no explanation, i.e., a distinct and determinate derivation of the possibility of a natural product that is possible in accordance with those two heterogeneous principles. (AA V: 412)

We would therefore need some other way of jointly pursuing the two maxims rather than pursuing one at the expense of the other. Kant’s own suggestion for such a heuristic is regarding the principle of mechanism as valid, but as subordinated to the principle of teleology:

For where ends are conceived as grounds of the possibility of certain things, there one must also assume means the laws of the operation of which do not **of themselves** need anything that presupposes an end, which can thus be mechanical yet still be a cause subordinated to intentional effects. Hence even in organic products of nature, but even more if, prodded to do so by their infinite multitude, we assume that intentionality in the connection of natural causes in accordance with particular laws is also (at least as a permissible hypothesis) the **universal principle** of the reflecting power of judgment for the whole of nature (the world), we can conceive a great and even universal connection of the mechanical laws with the teleological ones in the productions of nature, without confusing the principles for judging it with one another and putting one in the place of the other, because in a teleological judging of matter, even if the form which it assumes is judged as possible only in accord with an intention, still its nature, in accordance with mechanical laws, can also be subordinated as a means to that represented end[.] (AA V: 414)

This way of looking at things is essentially the traditional mechanical theist perspective, wherein nature is supposed to operate solely by means of the general principles of motion governing matter, and the hand of intentional design is seen in the specific structures and forms which happen to be the ones on which these general principles operate. Much like the artificer structures an artefact in such a way that, in it, the general principles of nature give rise to certain processes and activities rather than others, so God has structured nature such that the in itself blind and absolutely general principles of nature give rise to precisely those effects he intended. If one dismisses, along with many mechanists, the idea that unstructured nature can give rise, by means of the general principles of matter alone, to the specific structures of the world, then one needs to accept at least the initial structure of the universe as a brute given. This, of course, amounts to giving up on explaining the structure of nature naturally, but it does allow one to go on pursuing the mechanist programme without disregarding the structure of the universe. As we saw in chapters 1 and 2, this dilemma was central to Kant's thinking throughout the whole of his career. Here, we finally see a full response to the issue:

[S]ince the ground of this unifiability [of teleological and mechanist judging] 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, by which nothing is taken away from the teleological kind of explanation. (AA V: 414)

Thus, although it would be more adequate to understand nature through the concept of natural purposiveness, given that this concept is incomprehensible to us because of the peculiar constitution of our cognitive capacities, we need to understand it through the idea of intentionally designed artifacts in order to harmonize teleology and mechanism. The two principles are harmonized because we are only ever investigating the capacities of nature *given* these structural constraints in order to understand exactly how far these capacities extend. But to some extent these structural constraints will always themselves remain unexplained by the natural capacities. The fact that the structural component of a mechanist explanation is always itself mechanically unexplained is usually not much of an issue. In the case of natural purposes, however, it needs to be held in mind as essential, since they are entities for whom the fact that they have this rather than another structure is relevant in a way in which it does not seem to be for other mechanical interactions. By assuming that nature was disposed such that it give rise to these and precisely these structures, we are capable of regarding these structures as uncoincidental, but at the same time deny that something causally intervenes in nature, as characterized by its general principles alone, in order to bring

about these structures. The conclusion of the “Critique of the Teleological Power of Judgment” thus resonates with Kant’s perpetual hesitance between the adequacy of organicism and the heuristic commitment to mechanicism, clearly instantiated in the following comment from his 1766 *Dreams of a Spirit Seer*:

I am convinced that Stahl, who is disposed to explain animal processes in organic terms, was frequently closer to the truth than Hoffman or Boerhaave, to name but a few. These latter, ignoring immaterial forces, adhere to mechanical causes, and in so doing adopt a more philosophical method. This method, while sometimes failing of its mark, is generally succesful. It is also this method alone which is of use in science (AA II: 332)

This hesitance is the commitment to never bring the swing of Canguilhem’s pendulum, described in section 2.1., to a halt, however tempting it may be to proclaim that we have at last found our Newton of the blade of grass.<sup>22</sup>

## 4.4 Conclusion

In this lengthy chapter, I have tried to reconstruct and interpret the argumentive course of the tortuous and sinuous text that is the “Critique of the Teleological Power of Judgment”, or at least its first two parts.<sup>23</sup> Given the length and the complexity of this dialectical path, it is useful to recapitulate here what I take to be the most important conclusions of this chapter for the present dissertation.

First of all, Kant understands organisms as natural entities that inevitably suggest to us the idea of a unity of purpose. Such unity of purpose is to be understood as a functional or normative unity, namely unity of a rule or a norm for the species or the

---

<sup>22</sup> The scope of this dissertation prevents me from spelling out some of the most important results of this reading of Kant for the philosophy of biology. After all, my topic is transcendental epistemology, even though this requires, as I am continuously arguing, an eye on Kant’s perspectives on the life sciences. A study which is more focused on the philosophy of biology from similar premises is currently being undertaken by Joris Van Poucke, and I refer the reader to his work for further exploration of these issues.

<sup>23</sup> I have abstained from offering an interpretation of Kant’s subsequent treatment of physicotheology. This is not because I find it wholly irrelevant or uninteresting, but rather because I believe I am not yet well equipped do tackle this issue. The issue of physico-theology is intricately bound up with Kant’s transcendental idealist position on the relation between nature and normativity. In future work, I would like to reevaluate this position in the light of the findings of this dissertation, but I hope my reader shares my conviction that it should not be the subject of a summary treatment in the present work.

individual. Organisms are therefore to be judged as striving to maintain a certain unity of purpose, as under the obligation of answering to a certain norm. Such a norm, however, is not entirely static: it is plastic, since the organism can change its norm, i.e. change the relation and proportion of its relevant parts in order to adapt to contingencies.

In spite of the unity suggested by natural purposes, this unity should not be thought of as apart from, or over and above, the multiplicity of parts of which it consists. Kant rejects any interpretation of natural purposes as obeying to an externally imposed or transcendent norm, and insists that in order to maintain this, the whole itself cannot be distinct from the multiplicity of parts. The unity of a natural purpose is not, therefore, the unity of a transcendent principle of individuation or essence or idea, but rather that of the reciprocally productive and formative properties of the parts. In insisting on this, Kant reveals his distrust towards previous organismic theories for understressing the multiplicity of the organism to the detriment of a ruling part or transcending principle.

Despite his insistence on the concept of a natural purpose in this form, he insists that it is incomprehensible to us. Thus, the obligation to judge things as natural purposes embroils us in an antinomy due to the finite, i.e. discursive nature of our cognition, which means that the harmony between the universal and the particular is contingent for us, and always threatened by the possibility of failure. As a result, we cannot at the same time understand the unity of purposiveness and the naturalness of a natural purpose, since we understand a purpose in light of its irreducible particularity, and nature in light of its general principles and features. The antinomy of the teleological power of judgment thus reveals a division internal to our cognitive capacities itself.

In the Antinomy chapter of the “Critique of the Teleological Power of Judgment”, the antinomy is resolved in two steps. In the first step, Kant reveals that there is no straightforward contradiction between mechanism and teleology, since they are only opposed for our finite cognition. This allows us to think that there is a way in which they might harmonize, although they will never do so for us. But this leaves us with two contradictory demands of reason instead, and such a contradiction remains a danger to our capacity to coherently investigate nature. Kant shows that we can resolve this tension by heuristically adopting theistic mechanism, which regards organisms as intentionally structured for a certain end or purpose, and then investigates how the processes within these organisms and in which these organisms figure can be understood through mechanical means alone.

Those who still find resolution of the antinomy somewhat dissatisfying, may do so because Kant made no effort to overcome the finite nature of our cognitive capacities, and to protect us from the perpetual threat of cognitive failure. But this limitation is one that Kant himself would not regard as a failing, and instead regarded as a core aspect of his philosophy. In the next chapters, I will present an interpretation of this core aspect by first showing how reason itself is subject to disunity and therefore to an

antinomy even in its regulative demand for systematicity, and then how Kant's system of cognition is best understood not as a traditional unity, but as the kind of constantly threatened and adapting, reciprocally formative multiplicity that natural purposes are as well.



## Chapter 5 Systematicity and the Disunity of Reason

*We impose on one another, & it is but lost time to  
converse with you whose works are only [Aristotle's]  
Analytics  
Opposition is True Friendship  
- William Blake*

In this chapter, I will argue that Kant's conception of Reason and of Systematicity is such that a fundamental tension, and perhaps even an antinomy, can arise even within the heuristic use of this faculty. In 5.1, I will present a status quaestionis on the topic of transcendental systematicity, and introduce my own interpretation. Before developing this interpretation, I will first offer a sketch of the debate on systematicity and antisystematicity in the Enlightenment in 5.2. In 5.2.1, I will present the idea of systematicity as it was understood in Early Modernity, namely as a specific interpretation of Aristotle's semantics in terms of the tree of Porphyry. In 5.2.2, I will outline Locke's seminal criticism of that idea and indicate its influence on many important Enlightenment *philosophes*. In 5.2.3, I will then present Leibniz's attempt to overcome Locke's overly skeptical criticism with a new way of understanding and employing the idea of systematicity. I believe this background is important for a correct understanding of Kant's own conception and employment of this idea. In 5.3, I discuss the idea as it appears in the *Critique of Pure Reason*. First, I discuss the merely logical use of the idea of systematicity in 5.3.1, arguing that there too the commitment to completeness implies maximum particularity with maximal unity. In 5.3.2, I then discuss the transcendental use of systematicity, and argue that Kant saw a conflict arising between the distinct maxims to which it gives rise. In the *CPR*, the solution of this conflict is still the reminder that a merely reflective usage cannot run into an antinomy. In 5.4, however, I argue that the *Critique of Judgment* goes beyond this treatment by admitting that a reflective stance does not yet dispel all difficulties with the conflicting maxims. In 5.4.1, I present an interpretation of the introductions to the

third *Critique* as offering an answer to this issue, and argue that the principle of reflective judgment is meant to offer the solution. In 5.4.2, I then indicate how the antinomy of the teleological power of judgment can be better understood when it is read as a manifestation of the greater antinomy of systematicity. This latter discussion allows us to see the intimate connection between the tension between particularity and generality on the one hand, and that between sensibility and understanding on the other.

## 5.1 The Problem of Transcendental Systematicity

In the past few decades, Kant scholarship has been turning a more approving gaze at the peculiar section of the *Critique of Pure Reason* called the “appendix to the Dialectic of Pure Reason” (henceforth the Appendix). In that section, Kant concludes his critique of metaphysics with the claim that the ideas of reason do have a legitimate use, for even though we cannot take them as determinate metaphysical theses, or even as determinate transcendental principles, we can use them as guiding maxims in our quest for knowledge. This is because the ideas and maxims serve our striving for systematicity in knowledge, and such systematicity is required to speak of knowledge at all.

The thesis of the appendix raises some important questions of interpretation. We may ask, for instance, why this striving for systematicity would be necessary for knowledge, and why the forms of sensibility and the categories of the understanding are not sufficient as conditions for knowledge. After all, the point of the analytic is supposed to be that knowledge is possible, in spite of what the skeptic might want to claim, in virtue of the distinct contributions of sensibility and understanding. Now Kant seems to say that understanding and sensibility do not suffice to constitute knowledge.

This brings us to a second task, namely to show why the maxims of reason, despite being necessary conditions for knowledge, can still not be taken as determinate. In the analytic, Kant allegedly legitimated the constitutive use of the categories of the understanding by arguing that the categories are the conditions of the possibility of knowledge. But if this is true, then why can we not extend the same argument to the ideas of reason, i.e. why can we not conclude from their indispensability to knowledge that their constitutive use is legitimate?

To these two major issues a third needs to be added, namely the question why the striving for systematicity gives rise to the specific principles and ideas listed by Kant. The striving for systematicity is a very general goal, but Kant claims that several maxims and ideas can be derived from it and that these maxims and ideas are therefore indispensable guidelines in the quest for knowledge.



These puzzles explain why readers of Kant have often been extremely dissatisfied with the Appendix chapter. The first two problems, for instance, have been taken as indicators that Kant is blatantly contradicting himself here, for instance by Kemp Smith:

The teaching of this section is extremely self-contradictory, wavering between a subjective and an objective interpretation of the Ideas of Reason. The probable explanation is that Kant is here recasting older material, and leaves standing more of his earlier solutions than is consistent with his final conclusions. We can best approach the discussion by considering Kant's statements in A 645 - B 673 and in A 650 ff. - B 678 ff. They expound, though unfortunately in the briefest terms, a point of view which Idealism has since adopted as fundamental. Kant himself, very strangely, never develops its consequences at any great length. (Kemp Smith 2003: 547)

as well as by Walsh:

What this seems to say is that, unless ideas operated as they do, we could not make any judgments. It seems clear, however, that we could make judgments provided that we could conceptualise our data under the general restrictions of the categories; to prove that was the whole of the Analytic. Categories, and concepts generally, have a central role in Kant's scheme which makes them absolutely indispensable; without them there would be no experience at all. Kant cannot be saying that ideas too are necessary for the very possibility of experience, for this would undermine the whole distinction between ideas and categories. (Walsh 1975: 245)

Both authors object that the ideas and maxims of reason cannot at the same time be held to be necessary for knowledge and not amenable to constitutive use. Kemp Smith offers an explanation for this contradiction, namely that the appendix is an atavistic moment in Kant's thought, a remnant from his Leibnizian and Wolffian background. On such a reading, the contradiction would not be internal to Kantian philosophy, but rather between a more and a less mature state of that philosophy. Kemp Smith suggests that the Appendix should therefore be rejected as inconsistent with the mature transcendental philosophy. He also complains that Kant's brevity and inconsistency on the matter is unfortunate because this idea would become fundamental in the focus on unity in German (and British) Idealism.

Jonathan Bennett has equally protested against Kant's argument in the appendix:

A principle which was guaranteed to be permanently regulative would have to be necessarily incapable of (dis)confirmation. That privilege might belong to a principle which was somehow supposed by scientific endeavour as such, so that its confirmation would be question-begging and its disconfirmation self-refuting. I doubt if there are any such principles. (Bennett 1974: 275)

Bennett's objection is based on the question as to what kind of principle would be both necessary and non-constitutive in the manner Kant describes. But he indicates another issue at the same time, namely that of identifying such principles. The Ideas and maxims Kant identifies are supposed to be neither confirmable nor disconfirmable, yet still necessary. We might wonder why he took the ideas and maxims he proposed as satisfying this criterion.

Over the past twenty years, a thorough reappraisal of Kant's argument in the Appendix revealed that it is not nearly as problematic or straightforwardly contradictory as commentators like Kemp Smith, Walsh and Bennett presented it to be. Readings that mean to salvage Kant's claim that there are necessary though merely regulative requirements of knowledge stemming from the demand of systematization, could be called, with Ido Geiger (2003), transcendental interpretations. Several such transcendental interpretations have been offered by, for instance, Susan Neiman (1997), Nicholas Rescher (2000), Michelle Grier (2001), Ido Geiger (2003) and Henry Allison (2004).

What is remarkable about most such transcendental interpretations is that they seek to answer only the first two puzzles I identified at the start of this subsection, namely by offering a view on which systematization is a necessary condition for knowledge whilst being merely regulative, and not constitutive. The third puzzle, however, is deemed less important, for it is the demand of systematization, and not the maxims or ideas in which it expresses itself, which is of interest to these interpretations. Neiman (1997: 76-77) seems to suggest that Kant merely registered in them a variety of rules that scientists in fact use in their reasoning, which makes Kant's claim descriptive. Other readers, such as Grier (2001: 298-301), have seen little virtue in the concrete maxims and instead focused on the peculiar kinds of quasi-objects projected by the illusion of reason. Invariably, however, authors treat these maxims as mattering little, since they are all somehow equivalent to the general idea of systematic unity.

In this chapter, I will argue that this too summary treatment of the specific maxims and their relations has distorted our view both of Kant's argument for the necessity of systematicity to knowledge as of the nature of the systematicity he envisioned. By focusing only on the idea of systematicity itself, the striving for systematicity has been read as a striving for unity. This is unsurprising, given that Kant is usually read either in the context of Leibnizian philosophy, or in that of the German Idealism which he was to spawn. Many commentators, even those who defend the importance of the assumption of systematicity in Kant's philosophy (cf. Rescher 2002: 69 and Geiger 2003: 277), have suggested that this element is decidedly inherited from Leibniz, and that the Appendix account is some kind of leibnizianism made regulative rather than constitutive, heuristic rather than metaphysical. Reading Kant through the lense of German Idealism, however, leads us to focus on the feature of unity in the notion of systematicity. The

Appendix can then be read as an argument for assuming that nature is strongly unified, but at the same time denying that it is legitimate to assume this constitutively.

In this chapter, I will argue against this general reading by showing that unity is not Kant's sole concern in the project of systematicity, but that multiplicity and diversity are equally important. I will argue that Kant instead saw the problem of systematicity as an antinomy between two fundamental maxims: that of unity and that of specificity, and that he came to understand this problem through the 18<sup>th</sup> century debate on systematicity and antisystematicity erupting in the wake of Locke's *Essays*. On my reading, Kant is inspired by Leibniz's answer, but not in the way he is usually assumed to be. In the next section, I will present a brief outline of the debate on systematicity in Early Modernity. This will serve to understand the problem to which Kant is responding. In the final two sections of the chapter, I will offer my interpretation of Kant's position in the *Critique of Pure Reason* and the *Critique of the Power of Judgment* respectively. I will argue that the difference between these two accounts is less great than is traditionally assumed, and that in both the fundamental disunity of reason is the core topic. The principle of purposiveness as the principle for the reflective judgment is then to be regarded as Kant's response to the question of how to deal with this fundamental disunity, not his account of why we needn't worry about it at all.

## 5.2 The Spirit of Systems

### 5.2.1 The 17<sup>th</sup> Century Spirit of Systems<sup>1</sup>

The seventeenth century “modernists” profited themselves as the great innovators of philosophy and science. Our history since then has often taken them at their word, stressing their major departures from previous thought and hailing them for having liberated us from the great prejudices of the past, namely those stemming from Aristotelian and Theological conceptions of the world and of reality. In the previous chapter, I already indicated one way in which this is misleading, for the so-called banishment of final causes from natural philosophy is misunderstood if taken as a clear mark of secularization. In this subsection, I will discuss another way in which a

---

<sup>1</sup> This sections contains many passages, copied verbatim, from an earlier paper of mine on the topic of systematicity and antisystematicity in the Enlightenment. For a fuller treatment of the issue, I refer the reader to that text (Demarest 2013a).

fundamental critique of the Aristotelians' picture of science, harbours a peculiar continuity with it. This continuity is to be found in the idea that nature has a rational and systematic structure, and that hence we are warranted in examining it through certain rational means. In the next subsection, I will show how this assumption came to be attacked in various ways in the Enlightenment, predominantly the Enlightenment movements associated more closely with empiricism. In the final subsection of this section, I will present Leibniz's response to the challenge famously offered by Locke and adopted by many prominent *philosophes*.

When Aristotle speaks of science, he speaks of knowledge begotten through demonstration. Demonstrations yield knowledge because they are truth-preserving syllogisms that also preserve the necessity of their conclusions and that go from the prior to the posterior. In other words, demonstrations are syllogisms in which the three terms are related in a certain way. It is not my intention here to discuss the whole of the Aristotelian theory of knowledge and of science, but rather to indicate that, for many Aristotelians, demonstration is a matter of expanding and articulating the system of subordinate and superordinate concepts.

For many followers of the stagirite, this system of concepts formed what they called the tree of Porphyry, after the 3<sup>rd</sup> century neo-Platonic commentator of Aristotle. In Porphyry's text, the tree is more a vertical schema subordinating and superordinating certain concepts to each other. The rationale behind this schema is the following idea:

In each type of predication there are some most general items and again other most special items; and there are other items between the most general and the most special. Most general is that above which there will be no other superordinate species; and between the most general and the most special are other items which are at the same time both genera and species (but taken in relation now to one thing and now to another). (Porphyry 2003: 6)

Afterwards, however, the tree of Porphyry became associated not only with the idea of superordination and subordination, but also with the idea that all concepts might have a place in a tree diagram where the most general principles form the trunk and the most specific the branches. The tree diagram is generated by first taking the most general difference, and then either affirming or denying a specific difference of it. The affirmation yields one subordinate species, the negation another. Of each of the subordinate species, a further specific difference may then be either be affirmed or denied, yielding species subordinate to them. This process then goes on until we arrive at the lowest species, i.e. those concepts to which no others are subordinate, and under which only individuals are thought. All these "specifications" yield new "branches" of the tree, whose bifurcations connect the sturdy stem with the tender twigs.

The idea of superordination and subordination of concepts has a clear basis in some of Aristotle's explicit comments, although to my knowledge he never suggests that all

concepts can be organized into one hierarchical system of such a form. For Aristotle, the principles of a science are predications of concepts that are more universal, and thus superordinated. It is on the basis of the most general essential properties of the class of entities of which the science treats that we come to know the more specific properties. Science, for Aristotle, always moves from the more general to the more specific, deducing the properties of the specific from the general. In the same way, sciences themselves are subordinate to other sciences if the former treat of a species that is subordinate to the species of the latter.<sup>2</sup> Thus, scientific disciplines dealing with subordinate species are often regarded as applications of the higher-order sciences. Nevertheless, Aristotle insists (76a; Aristotle 1960: 67) that “the special principles of each science cannot be demonstrated; for the principles from which they would be demonstrable would be principles of all existing things, and the science of those principles would be supreme over all.” By this, he seems to mean that, at each (relevant) specification, there are some new properties which emerge and are the specific basic properties of that species. There can be no universal science from which all the specific properties of all the ontological kinds of things can be derived or deduced. This is why subaltern sciences are *distinct* from their superordinate sciences: they have their own principles and can use only their own principles. Nevertheless, a science might contain principles which are in fact provable or supposed by a superior science.

Throughout the Renaissance and the 17th Century, Philosophers targeted the philosophy of the schools by criticizing the traditional logic based on Aristotle’s *Organon*. The aspect of this doctrine that met with the most resistance was the so-called syllogistics. Francis Bacon, one of those figures who, in the *philosophes’* eyes, provided the great inspiration and instauration of modern science and its accompanying empirical mindset, formulated this criticism in the following way:

in the ordinary logic almost all the work is spent about the syllogism. Of induction the logicians seem hardly to have taken any serious thought, but they pass it by with a slight notice, and hasten on to the formulae of disputation. I on the contrary reject demonstration by syllogism, as acting too confusedly, and letting nature slip out of its hands. For although no one can doubt that things which agree in a middle term agree with one another (which is a proposition of mathematical certainty), yet it leaves an opening for deception; which is this. The syllogism consists of propositions; the propositions of words; and words are the tokens and signs of notions. Now if the very notions of the mind (which are as the

---

<sup>2</sup> The precise subordination of the various sciences is rather a complex issue in Aristotle’s *Analytica Posteriora*, which I will not disentangle here. The goal of the presentation here is only to indicate the general themes that were taken up by later authors, and that formed the framework of most subsequent thinking on science and knowledge.

soul of words and the basis of the whole structure) be improperly and overhastily abstracted from facts, vague, not sufficiently definite, faulty in short in many ways, the whole edifice tumbles. I therefore reject the syllogism; and not only as regards principles (for to principles themselves logicians do not apply it) but also as regards middle propositions; which, though obtainable no doubt by the syllogism, are, when so obtained, barren of works, remote from practice and altogether unavailable for the active department of the sciences. (Bacon 1875: 24)

In this passage, Bacon finds fault with the old *Organon* for its undue stress on the syllogism as the motor of reasoning, and therefore offers to replace it with his own new method, a new *Organum*. The typical Aristotelian demonstrative syllogism proceeds from the more general to the particular, thereby presupposing knowledge of the general, the principle. This is unacceptable to Bacon because it precludes any investigation into those principles themselves. In fact, his proposal, an *Organum* that is based on the method of induction, consists in an inversion of this process, a climbing of the scales of generality towards the most general:

hitherto the proceeding has been to fly at once from the sense and particulars up to the most general propositions, as certain fixed poles for the argument to turn upon, and from these to derive the rest by middle terms: a short way, no doubt, but precipitate; and one which will never lead to nature, though it offers an easy and ready way to disputation. Now my plan is to proceed regularly and gradually from one axiom to another, so that the most general are not reached till the last: but then when you do come to them you find them to be not empty notions, but well defined, and such as nature would really recognise as her first principles, and such as lie at the heart and marrow of things. (Bacon 1875: 25)

But an inversion of the order proper to Aristotelian logic leaves one thing unharmed: its basic conception of subsumption, of intension and extension. That Bacon still adhered to the hierarchical and unitary conception of nature as expressed in the trope of the tree of Porphyry, is clear from his exegesis of the myths surrounding the figure Pan in Greek literature in the *Sapientia Veterum*<sup>3</sup>. The key with which he tackles these tales, which he interprets as philosophical parables, namely the idea that “Pan, as the very word declares, represents the universal frame of things, or nature” (1878: 709), opens his path to the following observation:

---

<sup>3</sup> McRae (1957: 30-31) equally takes Bacon’s interpretation of the Pan-Myth as evidence that he subscribed to the hierarchical conception of nature, whereby the whole of nature is subsumed under one supreme principle. This is largely because, as Anderson (1948: 54) has pointed out, the *Sapientia Veterum* is an indirect attack on the ruling philosophies, via an interpretation of ancient myths as philosophical parables full of the wisdom of the empirically and materially minded way of mind from which Europeans had long since been led astray by scholasticism.

That the world is represented with horns, and that such horns are broad at bottom and narrow at top, has relation to the fact that the whole frame of nature rises to a point like a pyramid. For individuals are infinite: these are collected into species, which are themselves also very numerous, the species are gathered into genera, and these again into genera of a higher stage; till nature, contracting as it rises, seems to meet at last in one point. Nor need we wonder that Pan's horns touch heaven; since the summits, or universal forms, of nature do in a matter reach up to God; the passage from metaphysic to natural theology being ready and short. (1878: 710)

This assent to the hierarchical conception behind Aristotelian logic does not imply that Bacon is an Aristotelian: it is in fact likely that his major difference of opinion with Aristotle concerned precisely how we should conceive of the status of these “universal forms”, and how we should investigate them. But whatever Bacon may have meant with his concept of a form (and this is very hard to make out indeed<sup>4</sup>), he clearly believed that forms are related to each other after the logical manner in which Aristotelian species and genera are subordinated to each other. This is clear from his description of proper inquiry into forms in the second part of the *Novum Organum* by means of the example of heat. There, he suggests that heat is a specific kind of motion, distinguished from motion in general by a number of specific differences (1875: 151). In discussing the specific kinds of instances which may serve to guide one in proper inductive research, he states that “as Striking Instances lead easily to specific differences, so are Clandestine Instances the best guides to genera, that is, to those common natures, whereof the natures proposed are nothing more than particular cases” (1875: 160).

That Bacon leaves the scholastic conception of logical hierarchy unscathed is not to be mistaken for an indicator of the ultimately pre-modern mindset some believe to be behind his philosophical project. In fact, we can find the hierarchical notion of subsumption nearly everywhere in early modernity, not just in the writings of the Port-Royal rationalists, who merely saw the study of syllogism as rather unenlightening because they believed a great deal more errors of judgment result from relying on false or unclear principles of thought than from the inadequate employment of a rule of inference (Arnauld & Nicole 1992: 167), but also in those writings commonly taken to be precursors of the empirical line of thought that dominated the philosophes' theoretical enterprises. Gassendi's *Institutio Logica* includes an example clarified by means of a

---

<sup>4</sup> Some authors, such as Whitaker (1970), go as far as suggesting that Bacon's concept of a form is indefinite, and this because it harks back to a variety of concepts from other philosophies. For an overview of the literature's indecision with regard to Bacon's forms, as well as a recent attempt to remedy it, see Fletcher 2005. It seems likely, however, that Bacon's concept of a form anticipates Boyle's, which is decidedly non-Aristotelian, but may still allow for a proper classification (cf. Jones 2005).

schema situating the idea within the porphyrian tree (1727: 84), and Hobbes (1839: 30) equally believes that a proposition consists in a statement that one name comprehends another.<sup>5</sup>

Finally, the image of the tree of Porphyry also famously appears in the introduction to the French edition of Descartes's *Principes de la Philosophie*:

Thus the whole of philosophy is like a tree. The roots are metaphysics, the trunk is physics, and the branches emerging from the trunk are all the other sciences, which may be reduced to three principle ones, namely medicine, mechanics and morals. By 'morals' I understand the highest and most perfect moral system, which presupposes a complete knowledge of the other sciences and is the ultimate level of wisdom. (AT IXB: 14; Descartes 1985: 186)

This image of the tree is in no way original in Descartes, as was pointed out by Roger Ariew (1992). Ariew suggests that, in invoking this image, Descartes is deliberately using Aristotelian imagery to alleviate the remonstrances of scholastics, a hypothesis that is further supported by the fact that mathematics has no place in this tree. Nevertheless, there is no need to think that Descartes is wholly disingenuous here. For one, Descartes's claim that he agrees with Aristotle where the necessity for firm first principles in philosophy and science is concerned does not sound quite as hypocritical. This makes it plausible that he did indeed believe that the subaltern sciences are somehow legitimated by their deductive connection with the higher sciences. As a result, we should be skeptical towards the following claim made by Matthews, and identified by Ariew as the standard interpretation of Descartes arboreal metaphor:

Following upon some disturbing dreams in late 1619, Descartes's life work became the creation of a systematic philosophy which would encompass all branches of knowledge. The system would be based on a few undeniable principles, and all knowledge would be deduced from them, so that metaphysics, physics, mathematics, morals, and politics would all cohere. Knowledge is an organic whole, in which all fields have the same method. Descartes repeatedly used the metaphor of a tree: 'Thus philosophy as a whole is like a tree whose roots are metaphysics, whose trunk is physics, and whose branches, which issue from this trunk, are all the other sciences'. This doctrine of a single, all embracing method, is contrary to that of Aristotle, for whom the different fields of human knowledge all have their own subject matter and appropriate method. (Matthews 1989: 87-88)

---

<sup>5</sup> Hobbes does, of course, treat the hierarchical relation as a relation between names, and the relation between names as arbitrarily imposed. Unlike Bacon, he therefore need not regard nature as something that expresses the hierarchical systematic order, but he is committed to the fact that man's system of knowledge must take the form of a hierarchical system.



Matthews is wrong in believing that Descartes' invocation of the metaphor of a tree is a clear criticism of the Aristotelian division of the sciences. What is more plausible is that he is giving an account of the different interpretations of the tree metaphor in the two versions. As we saw above, Aristotle insisted that subaltern sciences are rooted in, and partly justified by, their superior sciences, but that they are nonetheless irreducible to those superior sciences. Matthews ascribes to Descartes the belief that such reduction is possible and warranted. Whether this is in fact the case, is not my concern here. I have solely wished to stress the wide spread of the systematic conception of knowledge in the 17<sup>th</sup> century, even among the great opponents of Aristotelian thought.

### 5.2.2 The Spirit of Systems under Attack

We have seen that, in spite of their disagreements with the scholastics, many heroes of Early Modern philosophy followed them in the belief that there is a systematicity to knowledge and to the universe, and therefore preserved the image of Porphyry's tree in some form or other. It is this lingering commitment to the systematic order that Locke sought to eradicate with his novel description of the inferential process, which did away with levels of generality and particularity altogether. Like his Early Modern predecessors, he believed that the focus on syllogism in reasoning was misplaced. His argument proceeds from the common 17<sup>th</sup> century theme that man has a natural light, which allows him to recognize validity or even truth and falsity without having to take explicit recourse to the highly artificial and complex syllogistics:

God has been more bountiful to Mankind than so. He has given them a Mind that can reason without being instructed in Methods of Syllogizing: The Understanding is not taught to reason by these Rules; it has a native faculty to perceive the Coherence, or Incoherence of its *Ideas*, and can range them right, without any such perplexing Repetitions. (Locke 1975: 671)

As is clear from this passage, Locke still believes that correct inference involves a good arrangement of the ideas, although he does not believe this arrangement needs to take the form of a syllogistic figure. Instead, he proposes another image of how terms are inferentially connected through intermediate ones:

To infer is nothing but by virtue of one Proposition laid down as true, to draw in another as true, *i.e.* to see or suppose such a connexion of the two *Ideas*, of the inferr'd Proposition. *v.g.* Let this be the Proposition laid down, *Men shall be punished in another World*, and from thence be inferred this other, *then Men can determine themselves*. The Question now is to know, whether the Mind has made the inference right or no; if it has made it by finding out the intermediate *Ideas*, and taking a view of the connexion of them, placed in a due order, it has proceeded rationally, and made a right Inference. If it has done it without such a View, it has

not so much made an Inference that will hold, or an Inference of right Reason, as shewn a willingness to have it be, or be taken for such. But in neither Case is it *Syllogism* that discovered those *Ideas*, or shewed the connexion of them, for they must be both found out, and the connexion everywhere perceived, before they can rationally be made use of in *Syllogism*: unless it can be said, that any *Idea* without considering that connexion it hath with the two other, whose Agreement should be shewn by it, will do well enough in a *Syllogism*, and may be taken at a venture for the *Medius Terminus*, to prove any Conclusion. But this no body will say, because it is by vertue of the perceived Agreement of the intermediate *Idea* with the Extremes, that the Extremes are concluded to agree, and therefore each intermediate *Idea* must be such, as in the whole Chain hath a visible connexion with those two it is placed between, or else thereby, the Conclusion cannot be inferr'd or drawn in; for wherever any Link of the Chain is loose, and without connexion, there the whole strength of it is lost, and it hath no force to infer or draw in any thing. (Locke 1975: 672-673)

Locke points out here that the very capacity to reason syllogistically is dependent upon the capacity to appreciate, besides the mode and the figure of the deduction, the connection between the terms thus related. After all, a syllogism fails to demonstrate if the terms are not in fact related in such a way as stated by the form of the propositions. In itself, this does not contradict Aristotelianism, for it also maintains that demonstration requires the terms to be related in the right way. The difference between the two thinkers is that Aristotelians believed that terms relate to each other in patterns of subordination, superordination and coordination within something like the tree of Porphyry, whereas Locke seems to require no such structure. This appears first of all from his shift in metaphor for the connection of terms, since he speaks of a *chain* of ideas rather than a tree of concepts. Unlike a tree, a chain suggests a linear rather than a hierarchical structure. Indeed, one of the major traits of Locke's philosophy is that he rejects the need for higher-order principles and maxims in reasoning as warrants for truth preservation and validity:

[S]ince the Knowledge of the Certainty of Principles, as well as other Truths, depends only upon the perception, we have, of the Agreement, or Disagreement, of our *Ideas*, *the way to improve our Knowledge*, is not, I am sure, blindly, and with an implicit Faith, to receive and swallow Principles; but is, I think, *to get and fix in our Minds clear, distinct, and complete Ideas*, as far as they are to be had, *and annex to them constant and proper Names*. And thus, perhaps, without any other Principles, but barely considering those *Ideas*, and by *comparing them one with another*, finding their Agreement, and Disagreement, and their several Relations and Habitues; we shall get more true and clear Knowledge, by the conduct of this one Rule, than by taking up Principles, and thereby putting our Minds into the disposal of others. (Locke 1975: 642-643)

It seems plausible to read this passage as Locke's charge against Descartes and the Cartesians, who maintained both the doctrine of clear and distinct ideas and that of general principles of thought. Locke seems to believe that, once we have accepted that there are clear and distinct ideas, and that the clarity and distinctness of an idea is an adequate philosophical ground, no further general principles are required as epistemological grounding. Hence, the whole machinery of different levels of generality, where reasonings of a more concrete nature are regarded as instantiations or applications of reasonings of an abstract nature, is redundant and, Locke adds, possibly harmful.

Locke has thereby radicalized the Early Modern critique of syllogistic reasoning in order to direct it against what he may have regarded as a residue of scholasticism: the systematic conception of reason and the world. One such form that this residue took is the persistent belief in general principles of reasoning. But Locke attacks another as well, namely the belief that nature forms a hierarchy of forms or kinds, a belief that still existed in Bacon, the patron saint of the Royal Society. In his chapter on General Terms, he disputes the authority of the traditional classification according to species and genera by arguing that instead of regarding a genus as an essence or form superordinated to the species, a lower-order essence, we should regard ideas as wholes consisting of other ideas, and the more general or abstract ideas as *parts* of the more specific ideas. According to Locke,

[t]his may shew us the reason *why, in the defining of words*, which is nothing but declaring their signification, *we make use of the Genus*, or next general Word that comprehends it. Which is not out of necessity, but only to save the labour of enumerating the several simple *Ideas*, which the next general Word, or *Genus*, stands for; or, perhaps, sometimes the shame of not being able to do it. But though defining by *Genus* and *Differentia* [...] I say, though defining by the *Genus* be the shortest way; yet, I think it may be doubted, whether it be the best. This, I am sure, it is not the only, and so not absolutely necessary. For Definition being nothing but making another understand by Words, that *Idea*, the term defined stands for, a definition is best made by enumerating those simple *Ideas* that are combined in the signification of the term Defined: and if instead of such an enumeration, Men have accustomed themselves to use the next general term, it has not been out of necessity, or for greater clearness; but for the quickness and dispatch sake. (Locke 1975: 414-415)

Locke here replaces the residual Aristotelian semantics behind logic with a modern one in terms of ideas. For Aristotelians, this semantics is one of concepts related to each other by the relation of intensional inclusion. Although Aristotle at times uses mereological language to characterize this relation, in doing so, as Aquinas put it, he “uses whole and part and comprehension equivocally”, (Aquinas 1954: 5). The genus is contained in the species in the sense that all specific differences listed in the definition

of the genus and every genus to which it is subordinate, are also listed in the definition of the species and every other species contained under the genus. But, in Aristotle's philosophy, there is no clear way in which one can say that the terms of a definition are parts of the definition in a *mereological* sense. To see this, note that every genus contains under it (at least) a pair of contradictory species, one of which the specific difference is affirmed, and one of which it is denied. This would mean that the genus is either a part of two distinct contradictory terms, or that the two contradictory terms overlap all but completely accept for the one specific difference. I doubt very much that Aristotle would be willing to accept either alternative, and would rather insist on the *sui generis* nature of the purportedly mereological relation of conceptual inclusion, i.e. used the mereological language merely metaphorically.

The Lockean picture is in that respect more easily understood as a mereological understanding of concept-relation. He starts from the ideas of particulars, which are complete, and which can be analyzed into their component ideas until each idea has been analyzed into simple ideas that cannot be further analyzed. The conceptually adequate definition of any idea is the list of simple ideas it contains. The idea of humanity is a component idea of Peter and Paul, but itself has components. It is therefore an abstractum. Now as to the question whether the same idea of humanity is contained in Peter and Paul, Locke can say that the idea of humanity in Peter is exactly *like* that contained in Paul because it consists of the same simple ideas. The question then becomes whether the same simple ideas are contained in an infinite possible amount of ideas. Here, the answer is offered by the fact that two simple ideas need not be token-identical, but can be type-identical, in the same way that two corpuscles in the corpuscularian philosophy can be: the simple idea red contained in my idea of the middle painting on the wall opposite me is then of the same type as the same simple idea red contained in my idea of the rightmost painting on that wall, but it is not the same idea.

Locke concludes from this that there is no need for universals, since universals are simply type-identical, and not token-identical, and hence that the logic of ideas allows us to be thoroughly nominalist on universals and essences:

It is plain, by what has been said, That *General and Universal*, belong not to the real existence of Things, but *are the Inventions and Creatures of the Understanding*, made by it for its own use, *and concern only Signs*, whether Words, or *Ideas*. Words are general, as has been said, when used, for Signs of general *Ideas*; and so are applicable indifferently to many particular Things: but universality belongs not to things themselves, which are all of them particular in their Existence, even those Words, and *Ideas*, which in their signification, are general. When therefore we quit Particulars, the Generals that rest are only Creatures of our own making, their general Nature being nothing but the Capacity they are put into by the Understanding, of signifying or representing many particulars. For the

signification they have, is nothing but a relation, that by the mind of Man is added to them. (Locke 1975: 414)

But Locke's position here only means that universals do not really exist as distinct individual entities, not that the classifications in terms of species and genera is thereby false. For it might be that, while there are no metaphysically distinct essences of species and genera, the particulars in the world are grouped according to clear-cut resemblance-classes. Locke realizes that this is a problem, since one of the major motivations for admitting the existence of universal essences (and not just particular ones) is that the regularities of nature suggest them. Specifically, the existence of universals is most suggested by the phenomena grouped under the concept of life that, namely the regularity of reproduction and the groupings of biological individuals into species or classes:

I would not here be thought to forget, much less to deny, that Nature in the Production of Things, makes several of them alike: there is nothing more obvious, especially in the Races of Animals, and all Things propagated by Seed. But yet, I think, we may say, the *sorting* of them under Names, is *the Workmanship of the Understanding, taking occasion from the similitude* it observes amongst them, to make abstract general *Ideas*, and set them up in the mind, with Names annexed to them, as Patterns, or Forms, (for in that sense the word Form has a very proper signification,) to which, as particular Things existing are found to agree, so they come to be of that Species, have that Denomination, or are put into that *Classis*. (Locke 1975: 415)

Locke responds to the taxonomist that there are indeed similarities in nature, some of which are very remarkable, but denies that this is evidence for the existence of species essences. His reasons for this are spread out over several chapters, but I will focus here on those arguments which challenge the biological evidence for species essentialism. The first challenge Locke issues is against the belief that the similarity between a parent and a child suggests a shared essence somehow responsible for that similarity:

Concerning the real Essences of corporeal Substances, (to mention those only,) there are, if I mistake not, two Opinions. The one is if those, who using the Word *Essence*, for they know not what, suppose a certain number of those Essences, according to which, all natural things are made, and wherein they do exactly every one of them partake, and so become of this or that *Species*. The other, and more rational Opinion, is of those, who look on all natural Things to have a real, but unknown Constitution of their insensible Parts, from which flow those sensible Qualities, which serve us to distinguish them from one another, according as we have Occasion to rank them into sorts, under common Denominations. The former of these Opinions, which supposes these *Essences*, as a certain number of Forms or Molds, wherein all natural Things, that exist, are cast, and do equally partake, has, I

imagine, very much perplexed the Knowledge of natural Things. The frequent Productions of Monsters, in all the Species of Animals, and of Changelings, and other strange Issues of humane Birth, carry with them difficulties, not possible to consist with this *Hypothesis*: Since it is as impossible, that two Things, partaking exactly of the same real *Essence*, should have different Properties, as that two Figures partaking in the same real *Essence* of a Circle, should have different Properties. (Locke 1975: 418)

Locke's move here is original in the sense that he takes traditional essentialism as a *hypothesis* which can then either square with, or be contradicted by, the phenomena. The hypothesis is that the species-similarity between parent and child is due to their both having the same species-essence. Locke invokes teratogeny as an empirical refutation of this hypothesis, for he takes birth defects to reveal that the child does not always maintain the species of the parent. Thus, he decides that is better to speak only of the individual essences constituted by the corpuscularian structure of a thing, and the possible similarities between these corpuscularian structures. In the next paragraph (1975: 419), he also cites nutrition as counter-evidence, since the materials of a creature of one species can be converted into that of another. This is problematic, Locke suggests, if species are indeed fixed and maintained.

These two arguments are directed at the idea that species are real because they explain the stability of these similarities. Together, they show that no such stability exists in nature. In a later chapter, Locke also adds an argument against the idea that species explain the clear grouping of natural entities (specifically animals and plants) by challenging the existence of such a clear grouping:

in all the visible corporeal World, we see no Chasms, or Gaps. All quite down from us, the descent is by easy steps, and a continued series of Things, that in each remove, differ very little one from the other. There are Fishes that have Wings, and are not Strangers to the airy Region: and there are some Birds, that are Inhabitants of the Water; whose Blood is cold as Fishes, and their Flesh so like in taste, that the scrupulous are allowed them on Fish-days. There are Animals so near of kin both to Birds and Beasts, that they are in the middle between both: Amphibious Animals link the Terrestrial and Aquatique together; Seals live at Land and at Sea, and Porpoises have the warm Blood and Entrails of a Hog, not to mention what is confidently reported of Mermaids, or Sea-men. There are some Brutes, that seem to have as much Knowledge and Reason, as some that are called Men: and the Animal and Vegetable Kingdoms, are so nearly join'd, that if you will take the lowest of one, and the highest of the other, there will scarce be perceived any great difference between them; and so on till we come to the lowest and most inorganic parts of matter, we shall find everywhere, that the several species are linked together, and differ but in almost insensible degrees. (Locke 1975: 446-447)

Locke invokes the idea of the Great Chain of Being, famously traced through the history of Western thought by Arthur Lovejoy (1966), according to which there are no gaps between the various beings in the world, because all entities form a chain of beings, ordered from the highest to the lowest. Here, Locke presents it as an empirical fact that frustrates all attempts to find clear general groupings according to taxonomical classes, and even a distinction between the living and the non-living itself. Furthermore, he insists that the distance between two adjacent links in the great chain are almost insensible. The result is a picture of nature as so diverse and so rich in forms that no rational system can neatly order it.

This line of thought in Locke became enormously influential in the Enlightenment. For many Enlightenment *philosophes*, the *Essay* revealed that our rationality or understanding is rather poor in its capacity, only being able to regurgitate received knowledge and dwell in a cycle of tautologies. Moreover, the systematic tendency of human rationality was a curse rather than a blessing, for in its hope to find nature mirror its own workings, it substitutes an impoverished, neat nature for the opulent though disorderly one that we know from our empirical inquiries. Denis Diderot phrased this idea as follows:

When we come to compare the infinite multitude of phenomena of nature with the limits of our understanding and the weakness of our organs, can we ever expect anything from the sluggishness of our labours, of their long and frequent interruptions, of the rarity of creative geniuses, than some shattered and scattered pieces of the great chain that links all things? (Diderot 1994: 562; my translation)

Diderot's friend and sometime collaborator Jean le Rond D'Alembert expressed a similar sentiment with regard to the lack of systematicity in natural products:

Of the truths that we have called isolated and floating, and that attach or do not seem to attach to any other, neither as consequence nor as principle, it is only in physics, and specifically in natural history that we can find examples. They consist above all in certain facts that experience discovers to us, and that seem, contrary to our expectations, to have no analogy at all with the facts that we constantly observe in the same species. For example, the sensitive quality in certain plants, or at least the apparent effects of this sensitive quality, a property that seems denied to all other plants, and limited almost uniquely to animated beings ; the multiplication of some animals without intercourse ; the reproduction of the legs of crayfish after they have been cut off ; the skill with which some animals, even some insects, seem gifted more than others; in a word, the peculiar properties that we observe in another genus of physical beings, and that seem contrary to those of other beings of the same genus. We can therefore define the isolated truths of which I am speaking here as particular truths that form, or seem to form, exceptions to general truths. (D'Alembert 1821: 135-137; my translation)

Like Locke, D'Alembert specifically singles out phenomena from the life sciences in order to reveal the vanity of our usage of general principles. Specifically, he shows that the phenomena of life at least appear to disobey the rigid taxonomical boundaries we construct. This blurring of kinds and sorts led Buffon to an anti-systematic approach to natural history, an approach with which Kant was intimately acquainted:

Since, without keeping ourselves to superficial knowledge the results of which can give us only incomplete ideas of the products and workings of nature, we want to penetrate deeper, and investigate with eyes that are at their most attentive to the form and the behaviour of her works, we are also surprised by the variety of design, as by the multiplicity of the manners of execution. The number of works of nature, although enormous, is responsible for only the smallest fraction of our astonishment; its mechanics, its art, its resources, even its disorders, merit our entire admiration. Too small for this immensity, overwhelmed by the number of wonders, the human spirit succumbs. It seems as if all that can be, is. The hand of the creator does not seem to have opened in order to give life to only a certain number of species; rather, he seems to have cast all at once a world of beings that are related and non-related, an infinity of combinations both harmonious and contrary, and a constancy of destructions and renewals. Such Idea of power this spectacle offers us! Such feeling of respect for our author this view of the universe inspires! What would happen if the feeble light that guides us would become clear enough to allow us to perceive the general order of things and the dependance of effects? But the vastest of spirits, the most powerful of geniuses would never elevate itself to such a height of knowledge: the first causes will remain for ever hidden to us, and the general results of these causes will be as difficult for us to know as the causes themselves. All that is possible for us, is to perceive some particular effects, compare them, combine them, and finally recognize an order that is more relative to our own nature, than proper to the existence of the things we are considering. (Buffon 1749a: 11-12; my translation)

At the zenith of Enlightenment *philosophie* in France, the idea that we will find in the world a structure that matches the cognitive faculties and demands, that does not constantly frustrate our attempts at understanding, and that can be measured by any remotely human standard of rationality, is completely absent. Focusing on the general, on the unity and on the systematicity of the universe will only bring us so much more sophistry and illusion. Instead, we are admonished to focus on the particularities in nature rather than its generalities, to train our eyes for detail rather than for outline, and to allow for disorder rather than demand order at any cost. This Age of Enlightenment was not so much the Age of Reason as the age of Reason's self-deprecation and self-castigation. In praising the humility of Locke's powers of the mind, the elusiveness of Newton's metaphysical commitment and the erudite ignorance of Bayle's Encyclopedism, the philosophes saw shadows where until only recently the



natural light of human understanding had shone bright, shadows they believed were cast by the flickering of this light itself.

### 5.2.3 The Spirit of Systems on the Defense

Although Locke's opinion would carry much authority on the European continent in the first half of the 18<sup>th</sup> century, it did not go unchallenged either. One of the most important challenges came from Leibniz, who attacked Locke's defamation of systematicity in his *Nouveaux Essais*. Leibniz's riposte is interesting because it is not merely a stubborn defense of systematicity against Locke's assumptions. In fact, the problem was a real one even internal to Leibniz's own system, for he shared Locke's focus on particular or individual essences and his insistence on the great chain of being and its concomitant idea of the plenitude of forms. This makes Leibniz's response all the more interesting, for it is an attempt to deal with the issue within the scope of a new philosophical atmosphere rather than an outright dismissal of newfangled systems. In this subsection, I will first discuss how Leibniz responds to the charge that there can be no generalities in nature because only particulars exist. Then, I will show how Leibniz is committed to the existence of a plenitude of logical forms. Finally, I will present Leibniz' solution to the problem through his optimism.

Logic was always a major interest of Leibniz', so much so that the two authors who arguably constitute the starting point of most 20<sup>th</sup> century Leibniz-scholarship, namely Bertrand Russell (1996) and Louis Couturat (1901) argued that it forms the basis and the key to his entire philosophy. While this claim has lost its attractiveness in light of the many studies into the relevance of other aspects of Leibniz's thought for his general philosophy (such as, but not exclusively, his physics, metaphysics, theology and biology), it has spawned serious inquiry into Leibniz's logic. One of the results of these investigative efforts is that we can no longer naively assume that logic remained unchanged between Aristotle and Frege. Another is that we have slowly acquired a better insight into what exactly sets Ancient logic apart from Modern (postfregean) logic. An important distinction is that Modern logic is concerned with propositions, where Ancient logic is concerned with concepts. The logically relevant components of a syllogism are not the propositions, but the terms related through the structure of the propositions and of the syllogism. A related distinction is that Modern logic is usually, and according to many ideally, extensional, whereas Ancient logic is intensional. Thus, Ancient logic is concerned with the modal relations between concepts rather than with the truth values of propositions or the set-theoretic relations between the extensions of concepts.

Like most Ancient logics, Leibniz's logic is concerned with the conceptual inclusion relation, i.e. with the specific entailment relation obtaining between the intensions of

concepts. Parkinson (1965: 14-15) has suggested that the key to understanding Leibniz's account of conceptual inclusion is his doctrine of an "alphabet of human thoughts". This doctrine is that all concepts ultimately consist of a combination of primitive and simple concepts. I believe that these simple concepts are what Leibniz elsewhere calls "perfections"<sup>6</sup>, defined as follows: "I call a perfection every simple quality which is positive and absolute, that is to say that expresses whatever it expresses without limits whatsoever" (G VII: 261; my translation). From this definition, several things follow: first, no perfection entails another, since no perfection entails anything but itself; and second, no two perfections contradict each other, since all perfections are positive. Hence, we can define every concept intensionally through the list of perfections that it contains.

There is a non-negligible relation between Locke's and Leibniz's respective accounts of the semantics behind their logic. Both reject the idea that there needs to be a specific order to the listing of the different perfections. This contrasts with the older hierarchical system which insisted on defining a concept through its immediately superordinate genus and a specific difference. Leibniz seems to agree with Locke that the idea of a specific difference as a somehow privileged or more proper predicate is unnecessary.<sup>7</sup> Whereas Locke defines concepts as abstract ideas consisting of a number of simple ideas, however, Leibniz defines them as intensions consisting of a number of simple concepts. This contrast both stresses the fundamental agreements and the fundamental disagreements between the two accounts.

The next step of Leibniz's logic to consider is his distinction between abstract and concrete concepts. In the *New Essays*, he gives his position in response to Locke's claim that abstract ideas are formed by abstracting from certain ideas contained in an idea of a particular:

I do not deny that abstractions are used in that way, but it involves an ascent from species to genera rather than from individuals to species. You see, paradoxical as it may seem, it is impossible for us to know individuals or to find any way of precisely determining the individuality of any thing except by keeping hold of the thing itself. For any set of circumstances could recur, with tiny differences which we would not take in; and place and time, far from being determinants by themselves, must themselves be determined by the things they contain. The most

---

<sup>6</sup> There may be some objections against the identification of perfections and simple concepts. I believe that wherever the two concepts don't seem to fully coincide, this is due to the fact that Leibniz is trying to have them do double duty as simple concepts and as divine attributes. In fact, Leibniz' stipulations work only for logical atoms, and not for divine attributes, which do not seem simple at all.

<sup>7</sup> Leibniz says, for instance, in the *New Essays* (G V: 270; Leibniz 1996: 291-292), that genus and differentia can often be switched (Cf. also Parkinson 1965: 15).

important point in this is that individuality involves infinity, and only someone who is capable of grasping the infinite could know the principle of individuation of a given thing. This arises from the influence – properly understood – that all the things in the universe have on one another. (G V: 270; Leibniz 1996: 289-290)

In this passage, Leibniz is criticizing Locke for thinking that we are ever in the possession of a complete concept of an individual, for a complete concept would be one which would allow us to grasp it and only it, in the way in which a Russellian definite description is supposed to do. According to Leibniz, we are only capable of individuating objects through demonstrative identification, because an individuating concept would be infinite in intension. This is of course due to his theory of individuation, which entails that from the essence or concept of an individual follow all seemingly contingent facts about it and about the world in which it exists. But I believe there is another reason: in order to have a complete concept of an individual, we would have to know of every perfection whether it is contained in the concept or not. Now, Leibniz arguably believed that there are an infinite number of perfections (Cf. Adams 1994: 148). Hence, a complete concept would either deny or affirm every perfection of the subject. An incomplete concept, however, can be finite, and hence can be grasped and used adequately by finite cognizers. Its downpart is, however, that through it we think a number of things, and the less complete our concept is, i.e. the less perfections it either affirms or denies of the subject, the less precise it is in scope, i.e. the more objects are thought under it.

An at first sight peculiar feature of Leibniz's criticism of Locke in the *New Essays* is that he agrees with the latter that there are only particulars, and that there are no such ontologically distinct things as universals. It would seem as if this commits him to nominalism. But Leibniz in fact argues that this does not follow, since Locke admits that our generalizations have some root in reality, namely the patterns of resemblances between different particulars. Leibniz agrees with this, saying that "generality consists in the resemblance of singular things to one another, and this resemblance is a reality" (G V: 271; 1996: 292), and then inquires, somewhat rhetorically, of Locke's spokesperson: "Then why not look for the essence of genera and species there too?" Leibniz seems to say that there is no need to accept the ontological existence of universals to admit that general concepts correspond with something in nature, for they may simply consist in the patterns of similarity that Locke believes to exist as well. This seems to allow for an intermediate position between nominalism and realism, which insists that universals are real without admitting that they are distinct ontological entities.

But this solution still lacks something, for the universals the existence of which Locke denies are not just factual patterns of similarity between existing particulars, but have some kind of modal import. That is to say, these generalities do not say anything about possibilities. This is not a problem for Leibniz, however, since he is not limited to

considering only actual objects. His solution to this is rooted in his conception of God. According to Leibniz,

all truths that concern possibles or essences and the impossibility of a thing or its necessity (that is, the impossibility of its contrary) rest on the principle of contradiction; all truths concerning contingent things or the existence of things, rest on the principle of perfection. Except for the existence of God alone, all existences are contingent. Moreover, the reason why some particular contingent thing exists, rather than others, should not be sought in its definition alone, but in a comparison with other things. For, since there are an infinity of possible things which, nevertheless, do not exist, the reason why these exist rather than those should not be sought in their definition (for the nonexistence would imply a contradiction, and those others would not be possible, contrary to our hypothesis), but from an extrinsic source, namely, from the fact that the ones that do exist are more perfect than the others. (Leibniz 1989: 19)

In this passage, Leibniz is making a distinction between the possible and the actual. All entities whose concepts do not involve a contradiction are possible, but not all of them really exist. Only some possibles are actualized, not all. This means that Leibniz could strictly speaking say that the modal overtones of universals rest on the considerations of the similarity relations between all *possible* objects, and not just all *actual* objects. But this raises the issue of the status of these non-actual possible objects: in what sense can their essences come under consideration at all? Leibniz responds to this in the *Monadology*: “God’s understanding is the realm of eternal truths or that of the ideas in which they depend; without him there would be nothing real in possibles, and not only would nothing exist, but also nothing would be possible” (G VI: 614; Leibniz 1989: 218). Thus, whereas actuals exist somehow<sup>8</sup> outside of the mind of God, possibles exist only in the mind of God. Theologically, it thus seems that actuality depends upon the divine attribute of omnipotence, whereas possibility depends upon divine omniscience.

In this way, Leibniz can show that both similarities between actual individuals and between possible individuals can be taken into consideration, and the latter give us the modal aspects of the universals. In the mind of God, which is the realm of logical intension, all possible objects are represented (all their essences are included in it). And it is here where Leibniz locates the Law of Continuity in logic: the possibles together form a continuous series of objects, each differing from the next only infinitesimally. I

---

<sup>8</sup> I insert this admission of vagueness here because it is not clear to commentators exactly in which sense actuals exist independently of the Mind of God. This is because despite, Leibniz’s constant polemicizing against Spinozism, it is at times unclear exactly how his position avoids collapse into Spinozism. I have no intention to enter that thorny debate here.

will here give a brief sketch of an argument which would have entitled Leibniz to this conclusion, but which was probably no more than somehow suspected by him.

First, we need to get clear on one of the implications of Leibniz's concept of a God. According to Leibniz, God is the *ens realissimum*, which implies that he has all the perfections. Leibniz tried to argue that the concept of an *ens realissimum* is consistent because no two perfections can contradict each other (since they are all positive and do not imply anything else). If we accept the reading that identifies the perfections with the simple predicates, it results that God has all of the infinite number of perfections, whereas each finite entity somehow lacks a perfection. Leibniz also argues that the perfections had by finite entities are derived from God: God is somehow the ground and the reserve of all possible predicates, and all finite beings are somehow privations of divinity.

If we accept that there is an infinite number of perfections, and that each perfection is either affirmed or denied, we can characterize the concept of each possible object through a series of 1s and 0s, where each place in the series represents a perfection, a 1 represents the affirmation of that perfection, and a 0 the privation of that perfection. God would be characterized by an infinite series of 1s, and nothing by an infinite series of 0s. Now we can prove that the series of possible objects is continuous. Assume an infinite list of possible combinations of 1s and 0s. According to Leibniz, each of those combinations yields a possible object (it is a complete concept, and no combination can result in a contradiction):

1	0	1	0	1	0	1	0	...
0	1	0	1	0	1	0	1	...
1	1	0	0	1	1	0	0	...
0	0	1	1	0	0	1	1	...
.								
.								
.								

We can now use Cantor's (1892) diagonal method to prove that this series of possible forms is of a cardinality greater than that of the series of natural numbers, and, on the condition that we accept the continuum hypothesis, that it has at least the cardinality of the continuum).<sup>9</sup>

---

<sup>9</sup> This means taking the first number of the first member of the list and switching from 1 to 0 or from 0 to 1, the second from the second member and switching from 1 to 0 or from 0 to 1, etc. The result will be an infinite series of 1s and 0s which represents a possible complete concept, but differs in at least 1 perfection from any other member of the list. Upon adding the resultant to the infinite series, we can again use the diagonal method to generate a new possible complete concept not already in the list, and can do so *ad infinitum*.

Now all of this is hypothetical, for first of all it is not clear whether Leibniz in fact believed that perfections work in the way assumed above, and it is all but certain that he did not anticipate Cantor's proof. But the fact that one of the central principles of Leibniz's metaphysics and Logic, namely the Law of the continuity of forms, follows directly from this position that he arguably had, of course adds some plausibility to the interpretation; for even if Leibniz was not well placed to *prove* that this series would be continuous, he could have easily suspected it.

We can now turn to Leibniz's answer to Locke's criticism of the idea of systematicity in nature. As could be expected from the comments above, this solution is that while the series of all *possible* objects constitutes a continuous series, the series of all *actual* objects contains discontinuities:

Able philosophers have addressed themselves to this question of whether there is a vacuum among forms, that is, whether there are possible species which do not actually exist, so that nature might seem to have overlooked them. I have reasons for believing that not all possible species are compossible in the universe, great as it is; not only with regard to things existing at the same time, but also with regard to the whole succession of things. My view, in other words, is that there must be species which never did and never will exist, since they are compatible with that succession of creatures God has chosen. But I believe that the universe contains everything that its perfect harmony could admit. (G V: 286; Leibniz 1996: 307)

This passage has a clear ring of irony to it: we find Leibniz, one of the great champions of the principle of continuity, whose dynamics was largely based on this principle, having to constrain it in order to guarantee the systematicity of the universe. In fact, this passage supports Nicholas Rescher's point that Leibniz's harmony is a real trade-off between maximal unity and maximal diversity:

One immediately striking feature of the Leibnizian standard of metaphysical perfection in terms of orderliness and variety is that this is a *conflict-admitting two-factor criterion*, and, as such, contrasts sharply with the long series of monolithic *summum bonum* theories that have so generally been in vogue in ethics – both before the time of Leibniz and afterwards, down to our own day. (Rescher 1981: 10)

Rescher's point is that God selects the best of all possible worlds not by means of a single hierarchy of goodness, but by a trade-off between a variety of conflicting desiderata. In this case, God is supposed to have balanced both the variety contained in the universe

with the unity or systematicity it exhibits.<sup>10</sup> This mirrors a trade-off between two of the divine attributes. In the passage quoted in 5.2.2, Buffon praised the incomprehensible variety of the universe as a sign of God's omnipotence. For Leibniz, this discredits God's omniscience, i.e. his supreme rationality. As supremely rational, God is not content to create a variety which frustrates all rational structure and coherency. If he is thereby limited in his creative capacity, it is of course only a limitation internal to his own nature and indicative of his *volition*.

In this way, Leibniz vindicates rationality by vindicating systematicity. The *philosophes* believed that our understanding and our rationality are poor instruments for grasping nature. Leibniz responds not by simply defending rationality's claims to hegemony, but by showing that, although an irrational world is *possible*, the actual world is selected at least partly for its rational structure. Of course, this does not give us a safe-conduct for projecting our current systems of knowledge onto the world, but it does warrant us to believe that, whatever the world may be like, it is rational in structure and hence comprehensible by our rational capacities.

## 5.3 The Disunity of Reason in the Critique of Pure Reason

### 5.3.1 Logical Systematicity

One of Kant's most peculiar claims in the Appendix to the Critique of Pure Reason is that the demand for systematic unity is a logical one. The claim is peculiar because it is not clear how these laws are logical ones – which is why Walsh (1975: 244) suggested that this claim bears witness to Kant's failure to distinguish between logic and methodology. The reasoning behind these objections is easily recognized: as logical principles, the maxims of reasons would indeed have the necessity Kant confers upon them, but is difficult to see how they are logical laws rather than mere methodological tools. In this subsection, I will argue that given Kant's own post-Leibnizian intensional logic, there is a good reason to assume that these principles are indeed logical ones.

---

<sup>10</sup> Rescher's position has been sharply criticized as an interpretation of Leibniz (e.g. Rutherford 1996), but I believe there are good reasons to assume that a reader like Kant, being aware of the aftermath of these arguments, would be sensitive to the conflict-admitting nature of the parameters of diversity and unity. In fact, I will argue, in sections 5.3 and 5.4, that Kant explicitly charged Leibniz with the disingenuousness of his own balance between the two unities.

The first clue Kant gives us is his thesis that, whereas the function of understanding somehow corresponds with judgment, that of reason corresponds with syllogism (CPR A 330 / B 386. This idea is a puzzle to any initial reading of the *Critique of Pure Reason*. Not only is it unclear how exactly reason is linked to syllogism, and how syllogism would lead to the idea of the unconditioned, it is unattractive to ally the function of reason this closely to a by now unattractive piece of Aristotelian logic.

The problem with this appraisal is that it misses the extent to which Kant changed the notion of a syllogism and its meaning for philosophy. In many cases where he seems to be talking about inferential rules, he is in fact attempting to characterize the principles behind reasoning through concepts alone, without the bearing of these concepts on reality, be it metaphysical or empirical. This is already apparent in a paper from the early 1760s, in which Kant argued that logicians had traditionally exaggerated the “subtlety” of the syllogistic figures, and argued against Aristotle that there was only one principle of the syllogism. Instead of listing correct and incorrect, reducible and irreducible syllogistic figures, Kant offered two general rules:

the first general rule of all affirmative syllogisms is this: A characteristic mark of a characteristic mark is a characteristic mark of the thing itself [...]. And the first general rule of all negative syllogisms is this: that which contradicts mark of a thing, contradicts the thing itself. (AA II: 49)

These principles amount to the following idea: the idea of conceptual inclusion is transitive, i.e. a concept contained in a concept contained in the definition of a thing is also contained in the definition of that thing. This means that in order to fully explicate a concept, we need to list all the concepts contained in it, not just immediately, but also mediately. Only then do we adequately grasp what a concept entails and what it excludes. The semantic rule on which the validity of syllogistic reasoning is based therefore entails that we have a grasp of the concepts we employ only when we have determined their inferential roles completely. In light of the discussion of 5.2.1, Kant's point here becomes clear, for he is stating that the concept of a syllogism is intimately bound up with the semantics behind Aristotelian logic, which is a semantics of a hierarchy of concepts related through the *sui generis* relation of conceptual inclusion. In this respect, Kant is explicating what was already at work in German rationalism, of which a late representative is Moses Mendelssohn, who in his 1785 *Morgenstunden* still wrote that all syllogistic reasoning depends on the correct analysis of concepts, and that this correct analysis is obtained by picturing the system of concepts as a tree, with the stem as the most general concept and the twigs as the most concrete concepts (2008-96).

A second aspect of Kant's theory points towards the principle of unity, namely that of the transcendental prototype. The transcendental prototype as an argument stems from an early attempt by Kant to offer a proof for the existence of God that bypassed his own famous objection that existence is not a predicate. In the 1763 treatise *The Only Possible*



*Argument*, Kant argued that the possibility of all things rested in the existence of a thing including, as a ground, all possible predicates (AA II: 78-79). I will not go into the details of the argument here, partly because I find the argument confused, and because Kant equally came to see it as confused. In fact, he came to believe that he was mistaken, in much the same way as Leibniz was, to attempt at deriving a proof of the existence of God from a logical requirement. As we saw in subsection 5.2.3., Leibniz believed that all perfections somehow reside in God, and that finite beings are to be understood as privations of God. In 1763, Kant adopted a similar argument, believing that it was in fact superior to the Anselman-Cartesian proof that Leibniz ultimately preferred. Kant's argument starts from the idea that there has to be a real ground for possibility. This move is obviously Leibnizian, because Leibniz too thought that possibility needs to reside somewhere, namely in the intellect of God. Through a series of reflections, Kant arrives at the idea that the real ground of possibility is an *ens realissimum*:

The data of all possibility must be found in the necessary being either as determinations of it, or as consequences which are given through the necessary being as the ultimate real ground. It is thus apparent that all reality is, in one way or another, embraced by the ultimate real ground. But precisely these determinations, in virtue of which this being is the ultimate ground of all possible reality, invest that being with the highest degree of all real properties which could ever inhere in a thing. Such a being is, therefore, the most real of all possible being, for all other beings are only possible through it alone. (AA II: 85)

This passage sounds very similar to the Leibnizian picture of the *ens realissimum* as that which has all the perfections. But the similarity immediately turns out to be more superficial than we would expect, for Kant continues as follows:

But this is not to be understood to mean that all possible reality is included among its determinations. This is a conceptual confusion which has been uncommonly prevalent until now. All realities are attributed indiscriminately as predicates to God or to the necessary being. That all these predicates can by no means co-exist together as determinations in a single subject is not noticed. The impenetrability of bodies, extension and such like, cannot be attributes of that which has understanding and will. Nor does it help if one seeks to evade the issue by maintaining that the quality in question is not regarded as true reality. The thrust of a body or the force of cohesion are, without doubt, something positive. Similarly, in the sensations of the mind, pain is never merely a deprivation. A confusion has seemingly justified such an idea. It is said: reality and reality never contradict each other, for both of them are true affirmations; as a consequence, they do not conflict with each other in the subject either. Now although I concede that there is no logical contradiction here, the real repugnancy is not thereby cancelled. Such a real repugnancy always occurs when something, as a ground, annihilates by means of a real opposition the consequence of something else. The

motive force of a body in one direction and an equally strong tendency in the opposite direction do not contradict each other. They are also really possible in one body at the same time. However, one motive force annihilates the real consequences of the other motive force; and since the consequences of each motive force by itself would otherwise be a real movement, the consequence of both together in one subject is nought. That is to say, the consequence of these opposed motive forces is rest. But rest is, indubitably, possible. From this it is also apparent that real opposition is something quite different from logical opposition or contradiction, for the result of the latter is absolutely impossible. (AA II: 85-86)

It is hard not to read this as a direct criticism of the Leibnizian interpretation of the argument, given what we have seen in 5.2.3. Leibniz argued that God had all the perfections as his properties or attributes, and that he could have all of them simultaneously because none of them can conflict with any other. Kant objects that all the properties can coexist in God as in a ground, but not as in a subject, i.e. not as properties of God, for two reasons. First of all, there are positive properties of other objects that cannot be properties of God, and secondly, and more importantly, whereas two realities cannot logically contradict each other, they can really cancel each other out.<sup>11</sup> Here, we see an early form of Kant's doctrine that we should distinguish between the logical possibility governing concepts and the real possibility governing objects. In the *Critique of Pure Reason*, Kant would eventually come to recognize that it is a philosophical category mistake to regard the ground of all possible predication as a possible object rather than as the horizon of the space of concepts:

Every *thing* [...], as to its possibility, further stands under the principle of *thoroughgoing determination*; according to which, among *all possible* predicates of *things*, insofar as they are compared with their opposites, one must apply to it. This does not rest merely on the principle of contradiction, for besides considering every thing in relation to two contradictorily conflicting predicates, it considers every thing further in relation to the *whole of possibility*, as the sum total of all predicates of things in general (CPR A 572 / B 600)

Kant calls this sum total of all predicates an ideal of reason, which is a transcendental substratum. He clarifies its role in judgment as follows:

Thus if the thoroughgoing determination in our reason is grounded on a transcendental substratum, which contains as it were the entire storehouse of material from which all possible predicates of things can be taken, then this substratum is nothing other than the idea of an All of reality (*omnitudo realitatis*).

---

<sup>11</sup> In 6.3.1, I will return to this argument about the distinction of the logical and the real.

All true negations are then nothing but limits, which they could not be called unless they were grounded in the unlimited (the All). (CPR A 575-576 / B 603-604)

Determination of a concept or an object consists in denying a part of this whole of reality of something:

the transcendental major premise for the thoroughgoing determination of all things, is none other than the representation of the sum total of all reality, a concept that comprehends all predicates as regards their transcendental content not merely **under itself**, but **within itself**; and the thoroughgoing determination of every thing rests on the limitation of this **All** of reality, in that some of it is ascribed to the thing and the rest excluded from it, which agrees with the “either/or” of the disjunctive major premise and the determination of the object through one of the members of this division in the minor premise. (CPR A 576-577 / B 604-605)

Kant insists here that the concept of the *omnitudo realitatis* is not primarily to be understood as the highest concept under which every other concept is subsumed, but rather as a whole *in* which all marks or determinations are included. This idea is clearly related to the idea of God as the *summum realitatis*, i.e. as that which has all the positive predicates. Nevertheless, Kant immediately criticizes the conflation between this whole of reality with the concept of God as most perfect being. The transcendental prototype is an ideal that we must assume in order to hope to completely determine concepts. At no point, however, can we believe that we are in possession of this concept, and even less that we than have a concept of an object that is the real ground of the possibility of things. With this criticism, Kant is recuperating his original idea of the sum total of possibility, but repudiating his earlier conviction that this yielded a basis for a proof of the existence of God.

In adopting this picture of the transcendental prototype, Kant is introducing a new metaphor for the relation obtaining between concepts, namely a *spatial* metaphor:

One can regard every concept as a point, which, as the standpoint of an observer, has its horizon, i.e., a multiplicity of things that can be represented and surveyed, as it were, from it. Within this horizon a multiplicity of points must be able to be given to infinity, each of which in turn has its narrower field of view; i.e. every species contains subspecies in accordance with the principle of specification, and the logical horizon consists of only smaller horizons (subspecies), but not of points that have no domain (individuals). But different horizons, i.e., genera, which are determined from just as many concepts, one can think as drawn out into a common horizon, which one can survey collectively from its middle point, which is the higher genus, until finally the highest genus is the universal and true horizon, determined from the standpoint of the highest concept and

comprehending all manifoldness, as genera, species, and subspecies, under itself.  
(CPR A 658-659 / B 686-687)

In this passage, Kant introduces an idea of logical space and suggests its analogy with the space of intuition. One of the reasons why the transcendental philosophy is so keen on asserting that space is a pure intuition and the form of outer intuition, is that it needs to be given before every spatial representation. Every space is conceived as a part of the one space, and hence relatable to any other part of space. Here, Kant argues that a similar situation occurs in logic: the reason why all our concepts can be brought to bear on each other is because they are all merely determinations of the one conceptual space. The logical principle of unity is precisely this: a concept is determined by its place in conceptual space, which means that it is determined by its relation to all other concepts. A concept that could not be related to other concepts, one that would be outside of conceptual space, could play no role in reasoning, and is hence not a concept at all. There are two other analogies between space and logical space: like space, logical space is not given in its totality, but rather as a horizon of all possibility, and like space, logical space is continuous, which offers a version of the law of continuity in logic.

What I have offered so far is the logical ground of unity, which comes down to Kant's logical holism. Now I will offer an interpretation of why Kant held the law of specificity to be equally required, namely by showing that it arguably follows from his intensionalism in logic. Kant's logic is determined by a central law that states that the intensions and the extensions are inversely related.<sup>12</sup> This means that, as more things are thought under a concept, less characteristic marks are thought in the concepts. At first, this principle seems rather obvious: the less restrictions we place on the membership of a class, the more members that class can have. However, in his attempt to reappraise intensional logic after the logical revolution of Frege and Russell, Clarence Irving Lewis (1918: 322) criticized the validity of this purported relation between intension and extension. His main objection is that there are a variety of cases where the coextensiveness of two concepts does not imply their identity of meaning, i.e. where the expansion of the meaning of a concept does not entail a restriction of its extension. For example, take a world containing forty objects, of which twenty are red cubes and twenty are green orbs. In this world, the class of green things and the class of orbs would be coextensive. But this does not mean that green and orb have the same meaning.<sup>13</sup>

---

<sup>12</sup> This law is best articulated in Kant's first proof for the fact that space is an intuition in the second edition version of the transcendental aesthetic (Cf. Allison 2004: 111).

<sup>13</sup> According to Nelson Goodman, the nominalist needs to accept that, in this case, green and orb are simply identical: "The nominalist's attitude stems in part, perhaps, from a conviction that entities differ only if their

Lewis' misgivings arise from a misunderstanding, however. When Leibniz and Kant speak of the extension of a concept, they mean by this its *possible* extension, not just in this reality, but in logic. Leibniz already argued against Locke that the modal claims we make about objects are not determined by the actual state of the world, but by the possibilities perceived in the mind of God. In Kant's de-theologized version of the argument, it means that there is a difference between the logical extension and the real extension of a concept, whereby the former is merely an extensional projection of its intension, and the latter a statement about the actual objects of experience. Part of the criticism Kant voices against his predecessors is that they mistook logical extension for an ontological category.<sup>14</sup> The significance of the law of specification in Kant's logic is then the following: for every modal nuance we can make, there needs to be a concept that grasps this nuance perfectly, and a set of objects that sets it apart from all other concepts. This law is a consequence of endorsing a logic that is concerned with intensions rather than extensions.

---

content at least partially differs. So far as individuals go, this is a truism; and any supposed exceptions, such as the case of two objects fashioned out of the same piece of clay at different times, clearly depend on the fallacy of ignoring the temporal or some other dimension. Further it is clear that two classes, however defined, are indistinguishable if they have the same members; classes are in a sense distinguished only by what is comprised within them. But the nominalist goes still a step further. If two distinct *entities whatever* have the same content, then a class (e.g. that of the counties of Utah) is different neither from the single individual (the whole state of Utah) that exactly contains its members nor from any other class (e.g. that of acres of Utah) whose members exactly exhaust this same whole. The Platonist may distinguish these entities by venturing into a new dimension of Pure Form, but the nominalist recognizes no distinction of entities without a distinction of content." (Goodman 1977: 36) Goodman thus claims that there can be no difference in meaning between sets or concepts that are coextensive (I will bracket for now the more difficult claim that the class does not differ from the individual that is the mereological fusion of all the members of that class, partly because Goodman's proposal to equate mereological parthood and membership is so problematic). In order to be able to say that, for instance, the state of Utah and the set of all the counties of Utah are not the same, we need to be able to say that some members of the set of the counties of Utah might not have been members of the state of Utah. For instance, some easternmost counties might through some legal peculiarity secede from the state of Utah and join the state of Colorado in order to be able to sell marijuana legally. In such a case, the set of all counties of Utah in the actual world and the set of all counties of Utah in the alternative (metaphysically, but not necessarily legally) possible world are not coextensive. Now, although it may be true at *any* world that the set of all counties of Utah and the set of all acres of Utah be coextensive, it is not true that the set of all counties of Utah is coextensive with the set of all acres of Utah across worlds. This shows us why extending from actual to possible extension allows us to express modal statements without recourse to extensions. However, this is precisely the move made by Leibniz, a move which installs a strict relation between intensions and extension such that they can be defined in terms of each other: the intensions can be defined in terms of the relations between possible objects, sets of objects or worlds, the relations between possible objects, sets of objects or worlds can be defined in terms of intensions. The distinction between Kripkean essentialism and Lewisian modal realism is partly a debate on which of these dual terms we take to be metaphysically more fundamental, and which as derived and reducible.

<sup>14</sup> Cf. 6.3.1.

In this subsection, I have tried to show why Kant could have regarded his maxims to be logical principles, and in so doing, I have already briefly indicated why they cannot be taken as straightforwardly metaphysical. In the next subsection, I will turn my attention to the concept of systematic unity in the Appendix in order to show that, when taken constitutively, these two laws contradict each other.

### 5.3.2 Transcendental Systematicity

In identifying the maxims of reason with logical principles, Kant can achieve two major steps in his argument. Firstly, he can argue that these principles are more than merely pragmatic aids in the search for knowledge, since they are imposed by the nature of reason and its proper business itself. These aims and ideas are not ones that we can adopt and discard at will: they are dependent on the structure of our cognitive faculty. Secondly, Kant can take his cue from logic to identify the transcendental principles, and can argue that there is a systematicity to them. Throughout his work, he relies on certain schemes, derived from his architectonic, to make claims about the necessity, indispensability and completeness of certain principles and concepts. These structures often mirror either the fourfold division of the categories, or the threefold division of the three moments of a single category. Just like Kant had claimed in the *Analytic* that the categories are the schematized versions of the forms of logical judgment, Kant here claims that the principles and ideas of reason can be regarded as the schematized versions of the principles of logical reasoning. As a result, the three maxims form a triad of thesis, antithesis and synthesis:

Reason thus prepares the field for the understanding 1. By a principle of **sameness of kind** in the manifold under higher genera, 2. By a principle of the **variety** of what is same in kind under lower species; and in order to complete the systematic unity it adds 3. Still another law of the **affinity** of all concepts, which offers a continuous transition from every species to every other through a graduated increase of varieties. We can call these the principles of the **homogeneity, specification** and **continuity** of forms. The last arises by uniting the first two, according as one has completed the systematic connection in the idea by ascending to higher genera, as well as descending to lower species; for then all manifolds are akin to one another, because they are all collectively descended, through every degree of extended determination from a single highest genus. (*CPR* A 657-658/ B 685-686)

Readers have tempted to focus almost exclusively on the first of the three principles, which dictates unity and homogeneity. This is undoubtedly because it is rather easy for 20<sup>th</sup> century readers to get an idea of what one would mean by a demand of unity through a Porphyry-tree of concepts. McFarland (1970: 20-21), for instance, does not

hesitate to read beyond what he sees Kant's awkward outdated logic towards the idea of reduction and unification prescribed by a hypothetico-deductive theory of science. On such a reading, Kant means to say that, in our investigation of the variety of phenomena given to us in experience, we should aim at subsuming them under ever more general laws or theories, despite the fact that we can have no guarantee that nature itself exhibits such unity. As a result, Kant is read as defending a progressivist theory of science, according to which we see ourselves as gradually approaching the asymptote of a grand unifying theory.

There are problems with this reading, however, and concomitantly with the recuperation of Kant to save a central but now unpopular aspect of logical positivist philosophy of science: the ideal of theory reduction and the progressivism that accompanies it. I want to focus on two related problems. The first is that it suggests that variety is a given, and reduction is an aim. It is easy to see that this is unsatisfying as a theory, because, as many readers have noted, there is no reason to assume that variety cannot be reduced in principle: whether a reduction succeeds is really an empirical affair. This objection is understandable, but it targets this reading of Kant rather than Kant himself. Indeed, Kant stresses that the variety of which he speaks is not an empirical given, but a logical and transcendental demand:

Also this law of specification cannot be borrowed from experience; for experience can make no such extensive disclosures. Empirical specification soon stops in distinguishing the manifold, unless through the already preceding transcendental law of specification as a principle of reason it is led to seek such disclosures and to keep on assuming them even when they do not immediately reveal themselves to the senses. (*CPR* A 657 / B 685)

I take Kant's meaning here to be the following: objects are given to us through concepts, and thought as objects through these concepts. For that reason, they are not given with their full, proper specificity. Throughout the Appendix, Kant denies that there are lowest species, i.e. concepts so specified that their extension necessarily comprises only one individual. Because we will never have the lowest species, we will forever need to think individuals through sortal terms, and thereby fail to grasp their individuating properties.<sup>15</sup> Nevertheless, in order to get a conceptual grasp on individuality, we need to go on specifying our concepts such that the specificity of individuals can be given to thought. In this passage, Kant is making a transcendental claim out of Leibniz's objection to Locke that we are never empirically presented with the full concept of an object:

---

<sup>15</sup> Cf. 6.3.1 for some alternative reasons why Kant took this to be the case.

Hence every **genus** requires different **species**, and these **subspecies**, and since none of the latter once again is ever without a sphere, (a domain as a *conceptus communis*), reason demands in its entire extension that no species be regarded as in itself the lowest; for since each species is always a concept that contains within itself only what is common to different things, this concept cannot be thoroughly determined, hence it cannot be related to an individual, consequently, it must at every time contain other concepts, i.e. subspecies, under itself. This law of specification could be expressed thus: *entium varietates non temere esse minuendas* (CPR A 655-656 / B 683-684)

Here, Kant is attacking a version of the myth of the given that states that the fullness of the empirical can be presented to thought as such, and that our conceptual activity is therefore supposed to grasp this empirical richness. Against this, Kant claims that the empirical never appears on the conceptual level as such, and has no authority over it. What forces us to take empirical richness into account is not that this richness directly impinges on our thought, but rather that we are constantly driven, by an internal demand of reason, to further specify our concepts, such that ever more fineness of grain in our conceptual grasp of the empirical can be achieved.

It is therefore wrong to read Kant as stating that we have a drive for unification that will be forever thwarted by the recalcitrance of experience. He is saying that the unrealizability of unification in the first sense is due to the fact that it is opposed by a second, conflicting maxim of reason, that of specification:

To the logical principle of genera which postulates identity there is opposed another, namely that of **species**; which needs manifoldness and variety in things despite their agreement under the same genus, and prescribes to the understanding that it be no less attentive to variety than to agreement. This principle (of discrimination, or of the faculty of distinguishing) severely limits the rashness of the first principle (of wit); and here reason shows two interests that conflict with each other: on the one side, an interest in the **domain** (universality) in regard to genera, on the other an interest in **content** (determinacy) in respect of the manifoldness of species; for in the first case the understanding thinks much **under** its concepts, while in the latter it thinks all the more **in them**. (CPR A 654 - 655 / B 682-683)

Kant is thus explicit that the two principles do not clearly harmonize, and are in fact in conflict with each other. The reason for this is, as the conclusion of the quoted passage indicates, that in the realm of logic, extension and intension are inversely related, and that generality therefore usually comes at the expense of information.

Kant thus denies that we can realize the unity which we seek because this unity is an internally conflicting demand. The systematicity of reason has an antinomical structure: it is guided by two mutually restricting principles. This reading is further suggested by the following passage:



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 satisfying this interest. (CPR A 666 / B 694)

This passage reveals that the single interest of reason, that of systematicity, invariably translates into conflicting methods and interests, and that for that reason it cannot be used as a constitutive idea. Interestingly, Kant goes on to indicate that the conflict of these principles has translated into two camps of natural philosophers, each stressing one of only two principles:

There is nothing here but the twofold interest of reason, where each party takes to heart one interest or the other, or affects to do so, hence either the maxim of the manifoldness of nature or that of the unity of nature; these maxims can of course be united, but as long as they are held to be objective insights, they occasion not only conflict but also hindrances that delay the discovery of the truth, until a means is found of uniting the disputed interests and satisfying reason about them. (CPR A 667-668 / B 695-696)

With my discussion from section 5.2. in mind, it is not hard to read Kant as addressing the high Enlightenment debate between proponents of the systematic approach, usually of a rationalist persuasion, and opponents of the spirit of systems, usually inspired by empiricist philosophies. Each party intends to do justice to one of the two demands of reason, but does so to the detriment of the other. Champions of unification gloss over the variety of nature, and make it ungraspable, whereas the panegyrists of variety lose hold of the intelligibility of the variety that they are witnessing. Furthermore, Kant realizes the central role played by the principle of continuity in this debate. He denies that we can claim that this continuity is an *observed, empirical fact* rather than a logical, transcendental or even metaphysical desideratum. Finally, the examples Kant uses indicate that he was most concerned with understanding the place of living systems in nature.

## 5.4 The Disunity of Reason in the Critique of the Power of Judgment

### 5.4.1 Reflective Judgment and its Principle

In the *Critique of the Power of Judgment*, Kant returns to the issue of the systematicity of knowledge already tackled in the *Appendix*. This raises the question of why he thought such a return was required, i.e. what failings in the original account he had identified in the meantime and how he sought to remediate them. According to many interpreters, Kant's revision is due to his discovery of the problem of contingency. Here is how this thesis is put in one of the best and most recent studies on the topic:

in the brief Appendix argument concerning the threat of diversity, Kant argues that only one of the regulative principles must hold of empirical nature, the principle of homogeneity. In the CJ, by contrast, immediately after Kant has presented the same threat of “infinite” empirical diversity (v:185), he glosses the principle of purposiveness, as it is to address this threat, as “the law of the specification of nature with regard to its empirical laws” (v:186), suggesting that all the regulative principles of systematicity, including those (specification and continuity) concerned with empirical diversity, are requisite for empirical knowledge, even of the most minimal kind. (Zuckert 2007: 44)

As my discussion of section 5.3. reveals, this common assumption is false, for Kant does require the principle of specification in the Appendix, and he already acknowledges that there is a tension between homogeneity and specification. Whatever the change that occurred between the first and the third Critique may be, it is not the discovery of a tension between the two principles. But then what can the additional discovery be? I would like to suggest that Kant had come to believe that his resolution of the antinomy of systematicity was insufficient. In the *Critique of Pure Reason*, Kant stated that no antinomy arose because the two principles are merely regulative, and not constitutive. But this is unsatisfactory in the same way that the initial solution to the antinomy of teleological judgment was unsatisfactory<sup>16</sup>, for we are still under possibly conflicting regulative demands. The *Critique of the Power of Judgment* is intended to overcome this additional difficulty and give us a principle for overcoming the conflict of the maxims of reason, namely the principle of purposiveness.

---

<sup>16</sup> Cf. Section 4.3.1.

In order to understand what the principle of purposiveness entails, we need to turn first to the formulation of the problem as it appears in the Introduction to the *Critique of the Power of Judgment*. Here, the problem is framed in terms of the contingency of the empirical laws with respect to the general order of nature required by the categories:

For it may certainly be thought that, in spite of all the uniformity of things in nature in accordance with the universal laws, without which the form of an experiential cognition in general would not obtain at all, the specific diversity of the empirical laws of nature together with their effects could nevertheless be so great that it would be impossible for our understanding to discover in them an order that we can grasp, to divide its products into genera and species in order to use the principles for the explanation and the understanding of one for the explanation and comprehension of the other as well, and to make an interconnected experience out of material that is for us so confused (strictly speaking, only infinitely manifold and not fitted for our power of comprehension).  
(CJ AA V: 185)

I take Kant to mean the following: The categories require that a certain coherence and unity is found in nature, which guarantees that causality obtains in it. But the fact that causal regularities obtain tells us nothing about what form these causal regularities will take. For all we know, some things or kinds of things, or all things or kinds of things, may exhibit causal features that cannot be related to other principles. For instance, it might turn out that a class of objects (which may be a unit class) behaves perfectly regularly and predictably in different circumstances, but a very similar class of objects behaves wildly different. Thus, the regularities exhibited by the former class give us no ground for inquiry into the latter class, and more importantly, we are always at risk that a tiny overlooked difference between two cases might mean that wholly different causal laws apply there. In other words, the principle of causality only requires that there are causal regularities in the world, but it does not require that such causal regularities are themselves related in a regular and orderly manner.

If we want to understand Kant's reason for thinking that this threat is brought about by the demand of systematicity itself, we need to see how the relation of causal laws relates to the relation between concepts. After all, Kant usually speaks of the particularity and generality of empirical laws and the particularity and generality of laws intermittently. The connection between the two is no longer obvious for us, but it was for anyone reflecting on the difference between mechanical philosophy and school philosophy. Kant did not think of laws as generalized statements of the form "all x are p" that are then somehow injected with modal robustness. Instead, he probably thought about laws along more Aristotelian lines as conceptual inclusion between concepts,

yielding a law more like “x-ness implies p-ness”. Such a statement needs no modal injection, since it is already intensional rather than extensional.<sup>17</sup>

Nevertheless, there are two major differences between how an Aristotelian and a mechanist would put these principles to work. For an Aristotelian, a substance has its specific causal powers by grace of being that specific kind of substance. Being of a different kind usually implies having different causal powers. For a mechanist, the critique of substantial forms implies two things: 1) a good causal explanation is given in terms of very general causal properties shared by all things, not in terms of the causal properties specific to this kind of thing, and 2) the most general properties shared by all things in terms of which we need to explain causally are non-sortal. Let me explain this further by a Kantian example that I have already briefly tackled in 1.1.3.

According to Kant, Newton improved upon Kepler by generalizing Kepler’s law. Kepler’s law states that all planets move in specific ways. In intensional terms, this implies that planethood implies a certain kind of behaviour, namely a specific movement pattern. Newton, however, shows that planets behave that way because all things with masses relate to each other in certain ways (through the law of gravitation). This law is not just a generalization, explaining why *planets* behave in a certain way because *all material objects* behave in a certain way; it also recasts the law from a statement in terms of sortals (planets, material objects...) in terms of a mass-term: mass. For Newton, according to Kant, planets behave in the way specified by Kepler’s law not because they *are* planets, nor because they *are* material objects, but because they *have* mass. Other such non-sortal concepts in terms of which a proper mechanist explanation can be cast are extension or properties sharing extension’s peculiarity of being continuous and quantitative.

This two-pronged reconceptualization of a proper causal principle explains why mechanicism incurred such problems with the issue of individuation. For an Aristotelian, the concepts in terms of which we can causally explain the behaviour of an entity are the same concepts that allow us to identify that entity as a specific (kind of) thing. For mechanicism, on the contrary, the causal principles in terms of which we can legitimately explain the behaviour of an entity should not be specific or individuating. For Kant, this problem was undoubtedly known at least because of the influence of Leibniz: after all, Leibniz complained that the mechanical philosophy required supplementation with some kind of substantial form to allow for individuation, or else collapse into Spinozism.

We can understand how this relates to Kant’s problem by reflecting on the following marginal note to the first, unpublished draft of the *Critique of the Power of Judgment*:

---

<sup>17</sup> Cf. footnote 18 and 19 of chapter 1 of this dissertation.

Could Linnaeus have hoped to outline a system of nature if he had had to worry that if he found a stone that he called granite, this might differ in its internal constitution from every other stone which nevertheless looked just like it, and all he could hope to find were always individual things, as it were isolated for the understanding, and never a class of them that could be brought under concepts of genus and species. (AA XX: 216)

Several things are to be noted in this passage. First of all, it clearly reveals that the issue of classification is at stake. Second, Linnaeus' problem would be exactly that signaled by Locke and his followers, like Buffon. Kant makes a distinction between the outward form and the inner constitution, which matches up with Locke's distinction between nominal and real essence. The nominal essence of a thing is the way we identify or classify it on the basis of our perceptions on it: two things share a nominal essence when we find that they resemble each other in some important respect. Locke's challenge to nominal essences is that we have no idea of telling whether the surface similarities we can perceive match up with the similarities between the inner structures of things. This inner structure is the mechanico-corpusecularian structure. Because of this possible gap, we need to be skeptical of our capacity to find the right similarities in nature, "right" here meaning: the similarities between the mechanically and causally relevant descriptions of things.

This allows us to restate the problem. For Kant, the problem is not that of classifying things, but of classifying things according to their causally relevant properties. The problem is double, as is revealed by Locke and Buffon. First of all, we cannot know whether the properties according to which we classify a thing match up with the properties in terms of which we should causally explain the behaviour of these things. Secondly, however, we do not even know whether the causally explanatory properties of things themselves allow for some classification. With this statement of the problem, we can better understand Kant's proposal for a solution in the form of a principle.

There are several versions of this principle in Kant's work. The first version in the published Introduction of the *Critique of the Power of Judgment* runs as follows:

that since universal laws of nature have their ground in our understanding, which prescribes them to nature (although only in accordance with the universal concept of it as nature), the particular empirical laws, in regard to that which is left undetermined in them by the former, must be considered in terms of the sort of unity they would have if an understanding (even if not ours) had likewise given them for the sake of our faculty of cognition, in order to make possible a system of experience in accordance with particular laws of nature. Not as if in this way such an understanding must really be assumed (for it is only the reflecting power of judgment for which this idea serves as a principle, for reflecting, not for determining); rather this faculty thereby gives a law only to itself, and not to nature. (CJ AA V: 180)

This principle basically states that we should regard nature as if it is rationally structured, with “rational” here meaning “corresponding to the demands and capacities of our rationality”. In other words, it is the principle of the purposivity of the structure of nature for our cognitive capacities. Nevertheless, it is only to be assumed as a rule for inquiry, not as a determination of nature as an object or a structure. In itself this is rather vague, but the reconstruction of the context I offered in this subsection allows us to make it more substantial.

I suggest that the principle of purposiveness of nature prescribes that the similarities and dissimilarities on the basis of which we form concepts ultimately match the similarities and dissimilarities between the causal properties of things. Another way to say this is that the principle prescribes that things (causally) behave similarly to the extent that they are similar, and that they (causally) behave dissimilarly to the extent that they are dissimilar. This principle establishes a harmony between the generality and the specificity in the form of the genus-species distinction. Remember that a species is defined by a genus, namely the more general features it shares with other concepts, and a specific difference. This specific difference does not overrule the genus, but rather *adds* a further specificity. According to Kant, the principle of purposiveness is therefore that the specificity that does occur in nature does not overrule the general features of nature.

This is precisely how the principle is formulated in the first Introduction to the *Critique of the Power of Judgment*: “Nature specifies its general laws into empirical ones, in accordance with the form of a logical system, in behalf of the power of judgment” (AA XX: 216). This formulation connects the doctrine more closely with that of the Appendix, because it demands that the conceptual structure of nature is isomorphous to that of logical space, and therefore conceivable:

The logical form of a system consists merely in the division of given general concepts (of the sort which that of a nature in general is here), by means of which one thinks the particular (here the empirical) with its variety as contained under the general, in accordance with a certain principle. To this there belongs, if one proceeds empirically and ascends from the particular to the general, a **classification** of the manifold, i.e., a comparison with each other of several classes, each of which stands under a determinate concept, and, if they are complete with regard to the common characteristic, their subsumption<sup>b</sup> under higher classes (genera), until one reaches the concept that contains the principle of the entire classification (and which constitutes the highest genus). If, on the contrary, one begins with the general concept, in order to descend to the particular through a complete division, then the action is called the **specification** of the manifold under a given concept, since the progression is from the highest genus to lower (subgenera or species) and from species to subspecies. This would be expressed more correctly if, instead of saying (as in common usage) that one must specify the particular which stands under a general concept, it were instead said that one

**specifies the general concept** by adducing the manifold under it. For the genus is (considered logically) as it were the matter, or the raw substratum, which nature works up into particular species and subspecies through various determinations, and thus it can be said, in analogy with the use of this word by jurists, when they speak of the specification of certain raw materials, that **nature specifies itself** in accordance with a certain principle (or the idea of a system). (AA XX: 215)

Two things are remarkable about this passage. First of all, it explicitates the problem of the antinomy in the Appendix. There, the antinomy arose because the principles that can be taken to hold for the realm of concepts insofar as it forms the semantics of logic, cannot be straightforwardly applied to nature, for antinomies could result. Here, Kant specifies his solution: specificity and generality can only run counter to each other if they do not end up forming a neat system. If we assume that nature forms such a system, we can regard both the quest for specificity and the quest for generality as compatible. In a way, this was already the solution of the first critique, but there it offered no concrete guidance of how to go about investigating according to such a picture. Here, we get the advice to hold things to be similar *to the extent that they are similar*, and dissimilar *to the extent that they are dissimilar*. Without the principle of systematicity, all we could get is the advice of holding things to be similar that are similar and dissimilar that are dissimilar. The latter advice would forbid any concept that does not completely cover a particular, and as we have seen, Kant believes *no* concept can ever fully cover a particular. In other words, the principle of reflective judgment forms the basis and justification of reasoning through analogies, i.e. to take the similarities between things as clues for expecting and investigating other similarities between them.

A second thing that is remarkable about the passage, and about the first introduction version of the principle, is that it is not phrased in the deistic terms of the published version. Here, the assumption is not that nature has been structured as a logical system by an understanding, but that nature structures *itself* as a logical system. Moreover, the metaphor Kant uses here is that of nature who herself species her matter, which is the general substratum, into specific forms. The attentive reader will recognize in this the picture of nature suggested by natural purposiveness, for there too Kant assumes natural purposes to be the entities that reveal the self-specifying, self-determining properties of the natural. There are two possible reasons why Kant may have left this out of his official introduction. The first is that this passage sounds profoundly atheist, since nature is conceived as herself having specifying, shaping powers, whereas in the official version she is conceived as systematically structured by an understanding. On such a reading, Kant may have been censoring himself in his official works. I believe this is only part of the story. As we have seen in subsection 4.3.3., Kant also resolves his antinomy of the teleological power of judgment by stating that nature might itself have the specifying properties required to engender natural purposes, but that we cannot

conceive of this given our cognitive limitations. As a result, we need to consider that specificity as intended and installed by a transcendent intentional agent. In the official introduction, Kant is stressing the latter aspect of the solution, whereas in the unpublished one he is stressing the former. This shift in focus may have been due to some degree of self-censorship, but the later position is not straightforwardly a retraction of the earlier.

In the next subsection, I will further develop the connection between the critique of teleological judgment and the antinomy of systematicity. Before doing so, however, I need to stress, for the sake of my discussion of the connection between the antinomy of systematicity and the critique of the aesthetic power of judgment in chapter 7, that the solution should not be overestimated, for the fact that the power of judgment employs this principle does not in itself guarantee that it will always be satisfied. On the contrary, although the principle guides us in *finding* the systematicity in nature, it does not make nature automatically *appear* systematized:

Now this transcendental concept of a purposiveness of nature is neither a concept of nature nor a concept of freedom, since it attributes nothing at all to the object (of nature), but rather only represents the unique way in which we must proceed in reflection on the objects of nature with the aim of a thoroughly interconnected experience, consequently it is a subjective principle (maxim) of the power of judgment; hence we are also delighted (strictly speaking, relieved of a need) when we encounter such a systematic unity among merely empirical laws, just as if it were a happy accident which happened to favor our aim, even though we necessarily had to assume that there is such a unity, yet without having been able to gain insight into it and to prove it. (AA V: 184)

Thus, although this principle is posited, we do feel amazement and delight if it turns out to be fruitful because we nevertheless realize that it is in no way legislative for nature or experience. This means that we also sometimes incur a feeling of frustration when nature proves recalcitrant to systematization, i.e. when our efforts at conceptualization of the particular are opposed by irreducible specificity. In chapter 7, I will discuss these pure feelings of delight and frustration and their relevance for understanding Kant's theory of constitution. Here, I would just like to signal some possible misunderstandings concerning the implication of this principle for Kant's theory of science, namely that it would somehow nevertheless be reductionist, unificatory and progressivist.

First of all, it does not follow from Kant's image of the genus-species relation that his image of nature is reductive, since in Aristotle, its originator, it forms the basis of an anti-reductionist picture of science. After all, nothing precludes that the specific differences of nature are themselves not reducible to higher-order principles. Moreover, as I will argue in the next subsection, Kant regarded mechanist theories as usually overly reductive, and criticized them for their insensitivity to the demand of specificity.



It makes more sense to think of the theory as unificatory, but here too there is an important difference between a reductive and a non-reductive version of unification. For a reductionist, there is a master science which is to form the backbone of unification, and which sets its conceptual standards as the standards of science in general. For a non-reductionist, the conceptual apparatus of the master science is just a scientific dialect with an army and a navy. In this picture, although the conceptual apparatus of one science is ultimately accountable to that of all others, there are no, or at least very little, preset rules of how differences are to be adjudicated. In a sense, unification requires that the concepts of all sciences need to be aligned to each other, and should not constitute insular idioms, but it does not require that all should speak a lingua franca, or that each idiom should simply renounce its specific expressive power.

Finally, although Kant's theory of science is importantly unificational, it should not be regarded as progressivist. To see this, I will show why three important possible readings of the progress of science in Kant cannot be upheld. First of all, one might think that Kant identifies the progress of science as finding ever more general genera and ever more specific species. But this is simply ludicrous, because it assumes that we have already stumbled upon adequate concepts, and all that is required is filling in the blanks in the conceptual system. On the contrary, we may reject whole systematizations because, upon closer examination, its concepts do not do justice to the similarities and dissimilarities they need to capture, i.e. prove to be inadequate.

But this presents us with a second way in which science can progress, namely by yielding ever more adequate systems of concepts. This is surely the drive behind Kantian science, but the concept of adequacy should not be misinterpreted. First of all, Kant cannot be held to believe that a concept is adequate if it corresponds with a conceptual juncture in the real world. As a transcendental idealist, he has no such hypostasized, metaphysically independent concept of nature at his disposal to function as the standard of adequacy. One could say here that Kant is an *empirical* realist, but the precise meaning of this doctrine is exactly what is under discussion. The empirically real, after all, is the product of constitutive activity, and the world the horizon of empirical reality, not a definite given whole with which we are faced in science. Thus, for transcendental idealism there is no nature in the sense of a transcendent standard of adequacy for our concepts.

Another transcendent standard can be imagined, namely the empirical. In this sense, science progresses as it finds ever more empirically adequate systems of concepts. But here we need to bear in mind that the empirical, for Kant, is only ever given through concepts. If this is true, then there is no concept-independent body of empirical data to which conceptual systems are held accountable. As our conceptual system changes, what is empirically given may change as well, since the empirical is now presented to the understanding by new concepts. In Kant's picture of science, the empirical changes along with the conceptual. This means that the empirical does not form an independent

standard against which rivalling conceptual systems can be tested and confronted. What *can* happen is that a change in the conceptualization at one area, prompted by the search for specificity and/or generality within the framework of a system of concepts, can lead to disruptions and incoherences in that conceptual system, because the system of similarities and dissimilarities no longer aligns properly. When this occurs, the obstacle might be overcome by a reconceptualization that may remain local, or may overthrow the whole previous conceptual system. The resulting novel system can then be said to be superior only because it overcame an obstacle that the former could not overcome. In this sense, the “superare” of “superiority” is not the overcoming of the old theory, but the overcoming of the obstacles, the disruptions, the challenges spawned by activities within the horizon of previous conceptualizations.

For Kant, I believe, adequacy is a property of the reciprocal alignment of concepts within a system, an alignment that is judged on the basis of whether it maintains the proportion of similarities and dissimilarities. Scientific investigation is the inquiry into the adequacy of conceptual systems to deal with these proportions. Although I cannot give that account here, I will offer the following conjecture on the correct interpretation of Kant’s theory of science: In science, we are interested in the adequacy of our conceptualizations. We test them by testing the alignment of similarities and dissimilarities. Newton does so by introducing a standard of judging the similar and the dissimilar in mechanics: the three laws jointly offer a framework of when we can speak of a difference or a change in motion (through absolute space, absolute time, and, most importantly, inertia). In experimentation, we set up our inquiry such that we receive an answer on what has remained stable and what has changed, what is similar and dissimilar. What is at stake here is not a “theory”, a “law”, a “principle”, but a *concept*, or rather how that concept relates to others. In sum, the standard of adequacy in science is a *structural* one.

#### **5.4.2 The Antinomy of the Teleological Power of Judgment and the Antinomy of Systematicity**

In this subsection, I will briefly revisit the antinomy of the teleological power of judgment in the light of the antinomy of systematicity. I will argue that the former antinomy is essentially a version of the latter, i.e. that the tension between mechanicism and teleology is a version of the tension between generality and specificity. In order to better grasp this, I will first return to a thesis from the 1763 *Only Possible Argument* with which I have already dealt in a different context, namely in 2.3.1, where I discussed the problem of mechanicism and teleology in Kant’s early reflections on embryology. There, Kant revealed the tension between two theses: the thesis that

the world is so constituted that the variety of its effects follows from a small set of general principles, and the thesis that certain effects of nature, namely, organisms, do not seem to be possible on the basis of but a small set of general principles. More concretely, he was picking up on the fact that the problem of living systems, and of generation in particular, was that it frustrated the mechanistic attempt to describe nature through a minimal set of general principles. In Kant's own time, two answers to this question were very popular.

The context of the discussion of mechanistic and teleology in the *Only Possible Argument* is the intention to show how the idea of a single intelligent creator can be relevant to our inquiries into nature. In 5.3.1, I briefly discussed Kant's proof for the existence of God from the *Only Possible Argument*. This proof makes up only the first part of the text, whereas the bulk is dedicated to the relevance of this thesis to various philosophical issues. One such result is that it allows us to regard nature as strongly unified, which suggests a mechanistic worldview in terms of general laws. The irony is that Kant goes on to show that the unity of nature seems disproved by the structure of living things, which seem too contingent and specific to be naturally generated. In light of this statement of the problem, Kant can be read as refusing metamorphosis and preformation because the former ends up denying the specificity of life, thereby violating the principle of specification, and the latter ends up denying the unity of nature, making each living being due to a specific creative act, thereby violating the principle of homogeneity. He disagreed with the traditional assessment of the desirability of preformation, arguing that it maintained the austerity of mechanistic only through rendering a variety of entities incomprehensible on natural principles. Leibniz's protestations that organisms were not supernaturally imposed because they were in nature from the beginning, is unconvincing because their natural inexplicability means that they are not truly integrated in nature. It seems that Kant learned from this discussion that accounting for life as natural does require you to violate your demand for austerity in principles. In reducing the number of natural principles, you are effectively losing a conceptual grasp over a wide variety of natural entities. Kant's own solution is to refuse the mechanistic basis of the answer and suggest epigenesis, which ascribes to natural entities specific productive and self-formative principles.

The very same logic seems to lie behind the antinomy of the teleological power of judgment. What is important about that antinomy is that it arises between two maxims, one being the thesis, the other the antithesis. In fact, an antinomy arises only because the two principles have the same standing as maxims of judgment. The antinomy would not arise if we were dealing with a demand for unity in explanation, embodied in the first maxim, and a set of phenomena recalcitrant to that maxim. As merely empirically given, organisms would not prompt us to restrain our desire for austerity. The antinomy arises only because we are led to regard organisms as not merely actually, but principally recalcitrant to mechanical explanation. However, they can only be

recognized as such if there is a maxim of reason demanding this. I submit, therefore, that the second maxim of the antinomy of teleological judgment is importantly related to the second maxim of systematicity, and that it is the demand for recognizing the specificity of living systems as natural purposes that constrains the principle of mechanical explanation.

In the previous chapter, I stressed that what Kant found mechanically inexplicable about natural purposes is that they exhibit a specificity for which we cannot account with the general principles of mechanics alone. I have argued that such general principles not only lack specificity, they also characteristically cannot function as a basis for individuation. We have seen, however, that Kant thinks of natural purposes as intrinsically specific and self-specifying and even self-individuating. Given his analysis of the conceptual requirements and resources of mechanicism, Kant cannot but conclude that mechanicism fails *in principle* to account for the specificity of life. But since reason has both a drive to generalize and a drive to acknowledge specificity, neither principle can simply be abandoned.

This also explains the connection with Kant's analysis of the limitations of our cognitive capacities. He claims that we are led to this antinomy because we cannot think the particular except through the general (cf. 4.3.3.). This means that there is always a tension between our need to understand the particularity of the objects which we try to conceptualize, and the fact that we only ever grasp this particularity through a general concept, i.e. through the general. For Kant, this was a direct result of the fact that we are divided between sensibility and understanding, and do not have an intuitive understanding.

## 5.5 Conclusion

In this chapter, I have argued that Kant's notion of systematicity was not uniquely one of unity and unification. On the contrary, the commitment to diversity is equally powerful and legitimate according to the critical philosophy. This, however, generates a tension, perhaps even an antinomy, between two possibly conflicting commitments. The Appendix to the Transcendental Dialectic of the First Critique and the Third Critique as a whole present solutions to this problem. Still, there is a disunity in reason itself, a disunity which may give rise to conflicts, and which corresponds with another, namely that between sensibility and understanding: the two commitments of systematicity are commitments to the irreducible contributions of two distinct faculties of the Gemüt, namely sensibility, which is responsible for particularity, and understanding, which is responsible for generality. In the next chapter, I will try to

develop this connection between the antinomy at the heart of systematicity and the multiplicity of the faculties of the mind, in an analysis of the *Gemüt* as an organic system of faculties.



## Chapter 6 The Disunity of the System of Faculties

*L'originalité de la doctrine des facultés chez Kant consiste en ceci: que leur forme supérieure ne les abstrait jamais de leur finitude humaine, pas plus qu'elle ne supprime leur différence de nature.*

- Gilles Deleuze

In this chapter, I will argue that Kant conceived of the system of faculties as an organic unity, i.e. as a normative unity of diverse parts that reciprocally produce and specify each other. In section 6.1, I present the broad lines of a dominant interpretative stance on Kant's doctrine of the faculties, specifically regarding the unity of multiplicity of the faculties. I focus on those interpretations that have sought to push onto Kant the view that there is a fundamental faculty in order to overcome the various Kantian dualisms. In the rest of the chapter, I argue against this view by first, in section 6.2, presenting an overview of the historical background to the debate before Kant. In 6.2.1, I discuss Christian Wolff's theory that the many different functions of the mind can be reduced to a single fundamental power of representation, the simplicity and uniqueness of which matches that of the substance in which it inheres, namely the soul. In 6.2.2, I present the theory of a major critic of Wolff's theory, Christian August Crusius, whose insistence on the plurality of fundamental forces influenced Kant. In 6.2.3, I present the views of Johannes Nikolaus Tetens on the subject, and linger on his idea of an epigenesis of the soul. Kant's reception of this idea will not be dealt with until chapter 7, where I will present my interpretation of the Kant's conception of the epigenesis of reason. With these historical backgrounds, both forward facing (6.1) and backward facing (6.2) in mind, I then turn to Kant's own views, in 6.3. In 6.3.1, I discuss the irreducibility of sensibility to understanding, and argue that it is central to Kant's transcendental idealism. Concretely, I argue that the specific contribution of sensibility is its intuitive nature, i.e. its capacity to present us with particulars, to allow us to refer. I then pass to a discussion of the underlying idea of the paralogisms, in 6.3.2, where I show that Kant

sought to dissociate the idea of the unity of thought from the kind of unity rationalist metaphysicians ascribe to the soul. In 6.3.3, I argue that although Kant saw some merit to attempt a unification of the faculties, and even regarded it as a necessary maxim of reason, he sincerely doubted whether it could be done, and even argued for the irreducibility of five distinct faculties. Finally, in 6.4, I argue that Kant's doctrine of the faculties can be better understood if we read him as regarding the unity of the system of faculties as an organic unity.

## 6.1 Against Kantian Dualism: the Hegelian and Heideggerian Legacies

The distinction and the mutual dependency of sensibility and understanding constitutes one of the core doctrines of Kant's critical philosophy, for it is the basis of the famous dictum that "thoughts without content are empty, intuitions without concepts are blind" (CPR A 51 / B 75). Indeed, it forms the core of Kant's theory of knowledge as presented in the *Critique of Pure Reason*:

these two faculties or capacities cannot exchange their functions. The understanding is not capable of intuiting anything, and the senses are not capable of thinking anything. Only from their unification can cognition arise. But on this account one must not mix up their roles, rather one has great cause to separate them carefully from each other and distinguish them. (CPR A 51-52 / B 75-76)

However exciting and central this thesis may have been, the history of the reception of Kant's philosophy has been marked by hesitations with regard to its correct interpretation and evaluation. Many readers have felt that this radical distinctness between the two faculties causes philosophical difficulties, and have therefore read Kant as envisaging the overcoming of this radical opposition. In defense of this reading, the following passage from the A-introduction to the *Critique of Pure Reason* is often cited:

All that seems necessary for an introduction or a preliminary is that there are two stems of human cognition, which may perhaps arise from a common but to us unknown root, namely sensibility and understanding, through the first of which objects are given to us, but through the second of which they are thought. (CPR A 15 / B 29)

For many post-Kantian readers, Kant here confesses his hopes that, one day, his own brute distinction could be overcome, and that the common root of the two faculties could be discovered.



The first elaboration of such a reading already occurred in the German Idealists, all of whom believed that Kant's philosophy was incomplete for being unsystematic. This lack of systematicity they identified in Kant's tendency to merely posit his oppositions and tables, whereas they demanded that all these posits be *deduced*. A proper deduction, they believed, would account for the different forms in terms of a unitary principle that gives rise to them rationally and/or genetically. Karl Leonhard Reinhold, for instance, believed that the resistance towards Kant's philosophy was due to problems with the *presentation* of that philosophy in the works of its originator, and that an emendation of presentation and argument was required (1795: 21-22). The most important such emendation, according to Reinhold, is the definitive treatment of the faculty of knowledge and its limitations. He intends to offer this through an analysis of what he calls "Vorstellungskraft", i.e. representative power, because he thinks of the concept of a representation as the genus to which the species of sensation, intuition, concept, idea, etc. belong (1795: 211). From this, he deduces that the various faculties responsible for these different species of representations ultimately belong to the one faculty of representation (1795: 212). In doing so, he effaces the Kantian distinction between the faculties to a certain extent.

Nevertheless, it was the German Idealists who most vehemently rejected Kant's distinction between sensibility and understanding. For Kant, it is essential that only intuition is intuitive, i.e. presents us with particulars, whereas understanding is always discursive, i.e. presents us with mere generalities. In the previous two chapters, we have seen how this peculiar constitution of our cognitive capacities results in clear limitations to our knowledge, and lures us into antinomies. For the German Idealists, these limits on our knowledge seemed arbitrary and unwelcome, since they frustrated any attempt to acquire real knowledge. This is why they sought to overcome the distinction and attempted to defend the idea of an intuitive understanding, which Kant rejected as impossible for us. Versions of this strategy are to be found in both Fichte and Schelling, but I will here focus on its most influential version, namely the one offered by Hegel.

It is in his early essay *Faith and Knowledge* from 1802 that Hegel presents his seminal criticism of the sensibility-understanding distinction in Kant. He first gives his general, Schelling-inspired criticism of Kant's philosophy as failing to complete the proper course of philosophy:

The Kantian philosophy remains entirely within the antithesis. It makes the identity of the opposites into the absolute terminus of philosophy, the pure boundary which is nothing but the negation of philosophy. We must not, by contrast, regard it as the problem of the true philosophy to resolve at that terminus the antitheses that are met with and formulated perchance as spirit and world, or soul and body, or self and nature, etc. (Hegel 1977: 67)

The criticism here is that the Kantian philosophy is characterized by a series of brute oppositions that it does not itself desire to overcome, but that true philosophy *should* overcome: “the sole Idea that has reality and true objectivity for philosophy, is the absolute suspendedness of the antithesis.” (Hegel 1977: 68) It does not take Hegel long to identify the foundation of this entire shortcoming in Kant’s theory:

The Kantian philosophy has the merit of being Idealism because it does show that neither the concept in isolation nor intuition in isolation is anything at all; that intuition by itself is blind and the concept by itself is empty; and that what is called experience, i.e., the finite identity of both in consciousness is not a rational cognition either. But the Kantian philosophy declares this finite cognition to be all that is possible. It turns this negative, abstractly idealistic side [of cognition] into that which is in itself, into the positive. It turns just this empty concept into absolute Reason, both theoretical and practical. In so doing, it falls back into absolute finitude and subjectivity, and the whole task and content of this philosophy is, not the cognition of the Absolute, but the cognition of this subjectivity. In other words, it is a critique of the cognitive faculties. (Hegel 1977: 68)

It is thus due to the brute opposition between sensibility and understanding that Kant remained stuck in an incomplete philosophy, incomplete in the sense that it could not arrive at, and could not want to arrive at, absolute knowledge. Given that it is an important feature of Hegel’s general argumentative strategy that he presents his own philosophy as emerging from the internal logic of the systems of his predecessors, he presents his own solution as somehow already anticipated by the very position it overcomes. In casu, Hegel sees Kant as already breaking down the rigid boundaries between the faculties as he attempts to solve his central problem:

How are synthetic judgments *a priori* possible? This problem expresses nothing else but the Idea that subject and predicate of the synthetic judgment are identical in the *a priori* way. That is to say, these heterogeneous elements, the subject which is the particular and in the form of being, and the predicate which is the universal and in the form of thought, are at the same time absolutely identical. It is Reason alone that is the possibility of this positing, for Reason is nothing else but the identity of heterogeneous elements of this kind. One can glimpse this Idea through the shallowness of the deduction of the categories. With respect to space and time one can glimpse it, too, though not where it should be, in the transcendental exposition of these forms, but later on, in the deduction of the categories, where the original synthetic unity of apperception finally comes to the fore. Here, the original synthetic unity of apperception is recognized also as the principle of the figurative synthesis, i.e. of the forms of intuition; space and time are themselves conceived as synthetic unities, and spontaneity, the absolute synthetic activity of the productive imagination, is conceived as the principle of

the very sensibility which was previously characterized only as receptivity. (Hegel 1977: 69-70)

In this passage, Hegel is arguing that, in the concrete course of his arguments, Kant cannot maintain his distinction between sensibility and understanding, because the activity of constitution requires that this distinction be overcome in the synthetic a priori judgment. Moreover, Hegel suggests that the centrality of the imagination and the doctrine of the figurative synthesis show that the forms of sensibility are themselves due to the power of spontaneity as well. I will not here develop a reading of the transcendental deduction, but I would like to point out that the distinction is never truly overcome. The figurative synthesis, for instance, does not *produce* the forms of sensibility and understanding; it rather seems to *act upon* these forms to transform them in some way. It is not contradictory for Kant to maintain that even the forms of sensibility must undergo some constitutive process in order to be presentable to the understanding in certain specific ways. On the contrary, we should expect that even the pure intuitions are in some way blind in abstraction from conceptual activity.

According to Hegel, however, even Kant must admit that there is an original synthetic unity in which all the faculties are grounded:

This original synthetic unity must be conceived, not as produced out of opposites, but as a truly necessary, absolute, original identity of opposites. As such, it is the principle both of productive imagination, which is the unity that is blind, i.e. immersed in the difference and not detaching itself from it; and of the intellect, which is the unity that posits the difference as identical but distinguishes itself from the different. This shows that the Kantian forms of intuition and the forms of thought cannot be kept apart at all as the particular, isolated faculties which they are usually represented as [sic]. One and the same synthetic unity – we have just now determined what this means here – is the principle of intuition and of the intellect. (Hegel 1977: 70)

Hegel thus concludes that the doctrine of the synthetic a priori itself requires that Kant abandon the distinction between the two faculties, and that he recognize their common ground and principle, which he believes to be the productive imagination (Hegel 1977: 71). This idea, erroneous as it seems when judged as an interpretation of Kant, was definitive for all subsequent scholarship, and in a way accepted uncritically, as was argued by Dieter Henrich (1994). It resurfaced both in the early and in the late 20<sup>th</sup> century, but never independently of the horizon formed by German Idealism.

The idea of the imagination as the common source of both sensibility and understanding also figures as a core thesis of Heidegger's admittedly eccentric Kant-interpretation, presented in his 1929 book *Kant und das Problem der Metaphysik*. In his account of the imagination, Heidegger focuses on its peculiar double identity. According to Kant, the lower faculties are characterized as intuitive and receptive, whereas the

higher faculties are discursive and spontaneous. The problem is that he characterizes imagination as intuitive yet spontaneous. This suggests that it cuts across the rigid distinction between the lower and the higher faculties. (Heidegger 1962: 136) Moreover, Heidegger notes that Kant admits the imagination to be a fundamental faculty, even though elsewhere he says there are only two fundamental sources of our mind (Heidegger 1962: 141-142). Slowly but steadily, Heidegger manoeuvres his way to the same conclusion suggested by Hegel:

The interpretation of the laying of the foundation of metaphysics has revealed that the transcendental imagination is not merely an external bond which fastens two extremities together. It is originally unifying, i.e., it is the specific faculty which forms the unity of the other two, which faculties themselves have an essential structural relation to it.

Is it possible that this originally unifying [bildende] center is that "unknown, common root" of both stems? Is it accidental that with the first introduction of the imagination Kant says that "we are scarcely ever conscious" of its existence? (Heidegger 1962: 144)

Both Heidegger and Hegel profile their reading of Kant as to be situated in the quest for grounding the possibility of knowledge. This suggests that they are dissatisfied with the answer Kant seems to give on first sight, namely that knowledge is possible through the specific cooperation of the faculties. Both seem to suggest that this overstates the finite nature of human knowledge and the inability to go beyond our subjective bounds. Instead, they claim that the unity of the faculties is necessary to account for a specific aspect of knowledge, namely its transcendence, even though both thinkers have very different conceptions of this unity.

In more recent times, namely in the last decade of the previous century, the issue of the faculties again came under the attention of philosophers outside of the often insulated community of scholars of German philosophy, namely in the writings of those thinkers usually grouped under the heading "Pittsburgh School". The context is best understood through the presentation of John McDowell's *Mind and World*.

McDowell devoted his series of John Locke lectures to the peculiar nexus of epistemology and philosophy of mind that became so problematic after Wilfrid Sellars' *Empiricism and the Philosophy of Mind* and Donald Davidson's "On the very Idea of a Conceptual Scheme". McDowell read the arguments of these two seminal essays as converging on a very similar conclusion, namely the inadequacy of empiricism. Sellars' attack on the Myth of the Given revealed that the kind of brute givens philosophers have taken empirical data to be cannot play the foundational role in epistemology often accorded to them. This foundational role was taken to be warranted by the fact that givens are somehow immediate, whereas the constructions built upon them are irremediably mediated. According to Sellars, this idea of immediacy is untenable, a

conclusion he partly reached through a philosophical reenactment of Kantian and Hegelian arguments.<sup>1</sup> Donald Davidson, in his turn, argued against what he perceived as residual empiricism in Quine's philosophy by identifying a third dogma of empiricism and subsequently criticizing it. This dogma is that concepts or conceptual schemes stand over and against something external to them which contributes their *content*. Once this dogma is abandoned, Davidson argues, we have ipso facto abandoned the very idea of *empiricism* as well (2001: 189).

McDowell takes up Kant's doctrine of the relation between conceptual schemes and the given as captured in his famous dictum that "thoughts without content are empty, intuitions without concepts are blind" as a basis for the solution of the problem raised by Sellars and McDowell. This problem, it seems to me, is actually an interrelation of two problems, one from epistemology and one from the philosophy of mind. Given McDowell's sources, this is understandable, since both Sellars and Davidson usually treat these two domains as entangled. But in McDowell, this connection also leads to a bivalent evaluation of Kant's own doctrine, bivalent in the sense that he regards it as good in dealing with the first problem, and bad in dealing with the second.

The first problem is described by McDowell as *an antinomy*: "experience both must and cannot stand in judgement over our attempts to make up our minds about how things are" (McDowell 1994: xii). More specifically, the antinomy arises because, in absence of anything that plays the role of the given, there is no external constraint on thought, thereby making our reasoning about the world a mere game. This is a problem given that we usually think knowledge should have a justification, preferably at least in part in an evidentiary basis. But Sellars and Davidson have argued that the given cannot play that particular role of external constraint precisely because of its externality. The solution to this antinomy would be to think of the constraint as already arising within the space of concepts, rather than as arising on its edges and originating outside of it.

The other problem McDowell attempts to tackle is already announced in his formulation of the antinomy quoted above: he does not just want thought to be constrained in a non-trivial way, he wants it to be constrained by *how things are*. In other words, he wants thought to be meaningfully about an external world. The motivation behind this is that the Davidsonian and Sellarsian focus on the distinction between the natural and the normative frustrates the naturalist commitments of twentieth-century philosophers, for it becomes unclear how normatively engaged beings can become engaged with the non-normative structure of the world at all. Here the problem is that,

---

<sup>1</sup> Sellars (1997: 45) notoriously describes this work of his as "incipient Hegelian Meditations", and, in doing so, seems to contrast it explicitly with Husserl's Cartesian meditations. Where Husserl's work is meant to show us, amongst others, the contemporary value of the idea of immediate presence in the form of evidence, Sellars joins continental critics of the metaphysics of presence in rejecting the alleged grounding nature of presence.

due to the radical difference between the nature of our mind and the nature of the world, the two could never meet. McDowell's own answer to this problem is a Platonized naturalism.

This two-pronged problem forms the background of the following evaluation of Kant's philosophy: "Does Kant credit receptivity with a separable contribution to its cooperation with spontaneity? It seems that the answers "No" and "Yes are both applicable" (McDowell 1994: 41):

From the standpoint of experience, the answer is "No". If one posits an empirically separable contribution from receptivity, one commits oneself to something Given in experience that could constitute the ultimate extra-conceptual grounding for everything conceptual, and it is a way of putting a central Kantian thought to rest that that idea must be rejected. For Kant, experience does not take in ultimate grounds that we could appeal to by pointing outside the sphere of thinkable content. In experience we take in, through impacts on the senses, elements in a reality that is precisely not outside the sphere of thinkable content. (McDowell 1994: 41)

Here, McDowell praises Kant's solution to the problem, although he immediately also raises the questions that a good account of Kant's problem needs to answer, namely, why sensibility should be distinguished at all on this picture and, if it is distinguished, how it comes to be so marvelously integrated. In this chapter, I will suggest answers to these two problems by indicating how Kant thinks sensibility constrains thought and how it comes to acquire the normative force to do so.

McDowell also rejects Kant's solution in another sense, however:

But Kant also has a transcendental story, and in the transcendental perspective there does seem to be an isolable contribution from receptivity. In the transcendental perspective, receptivity figures as a susceptibility to the impact of supersensible reality, a reality that is supposed to be independent of our conceptual activity in a stronger sense than any that fits the ordinary empirical world. (McDowell 1994: 41)

McDowell is thus criticizing Kant for his confusion of reasons and causes in the transcendental perspective, even though he strictly observed, and was even largely responsible for the recognition of their distinction at the empirical level. Moreover, McDowell suggests that it is, paradoxically, only by accepting this notion of a supersensible inaccessible reality that Kant could maintain his idealism, whereas the abolishment of the very idea of the supersensible would allow us to embrace empirical realism without the constant reminder that it is at the same time a transcendental idealism. McDowell credits Hegel with the necessary step for such a completion, agreeing with the latter that absolute idealism is precisely what brings us into contact with the world in itself once again. (McDowell 1994: 44) The fact that two other major

figures of the movement to which McDowell belongs, namely Robert Brandom and John Haugeland, choose to build on Hegel and Heidegger respectively is telling in this respect.

This dissertation is not an interpretation and defense of transcendental idealism. Nor could it be, given the immensity of the task of defending this much-assailed doctrine. Nevertheless, I do want to signal a peculiarity in McDowell's criticism, namely his interpretation of transcendental idealism and its implications for the interpretation of the relation between transcendental idealism and the doctrine of the faculties. His interpretation of transcendental idealism is built around the so-called doctrine of affection, namely the doctrine that the empirical element of our thought is due to the causal influence of some external world. Since the very publication of Kant's work, readers have been debating over the coherence and meaning of this doctrine. A traditional criticism is the following: Kant believes us to be causally affected by the supersensible in experience, even though he believes causal processes cannot be ascribed to the supersensible, since causality is a category for our empirical thought, and not for transcendental reality. I would like to point out that the problem is a constructed problem, because *Kant nowhere advances the doctrine* – it is only ascribed to him on the basis of some terminological peculiarities and his readers' expectations of what he should or could hold. The very fact that any possible reading leads to incoherences might explain why Kant abstained from giving any doctrine, since no such doctrine can be given. The desire to have an account of the origin of the empirical is a metaphysical demand, and Kant might be consistent precisely in not meeting such a demand. The whole issue turns on the question of whether we may interpret the supersensible metaphysically.

McDowell obviously ascribes to the traditional reading of the doctrine of affection, and thereby reads Kant as saying that transcendental idealism is true because there is a supersensible reality which we cannot come to know. Yet, he believes this doctrine to be fortuitously independent from the intuition-concepts story. Mind that Hegel and Heidegger were under no such illusion. Hegel clearly realized that transcendental idealism was more due to the distinction between sensibility and understanding than due to the nature of the thing-in-itself. I believe he was right in this: transcendental idealism is more a doctrine about the incompleteness of the higher faculties and their resultant inability to yield knowledge by themselves than about the metaphysical status of the supersensible. This is why McDowell might be underestimating the contributions of sensibility to knowledge, even though he is right to dismiss the idea that it contributes foundational bare givens. In subsection 6.3.1, I will argue that, according to Kant, sensibility is not just responsible for the empirical constraints on thought (it is receptive), it is also responsible for the capacity to think particulars (it is intuitive), and that it is the latter capacity which constitutes its distinct contribution.

In this section, I have sketched the background against which an interpretation of the doctrine of faculties may be philosophically relevant, and not just an issue of philosophical antiquarianism. But this does not waver historical considerations; on the contrary, in order to better understand Kant's account, we cannot just read him in light of the debates he was to spawn – we must also read him in light of the debates in which he himself sought to intervene.

## 6.2 Faculty Psychology in German Rationalism

### 6.2.1 Wolff and the Substantial Unity of the Representative Power

Arguably the most important source and target of debates on faculty psychology in 18<sup>th</sup> century German *Schulphilosophie* is Christian Wolff's theory of the soul. It is well known that Wolff was a seminal figure in German philosophy in the 18<sup>th</sup> Century, and that before the advent of Kant's mature philosophy his thought dominated academic debates in the universities of the *Vielstaaterei*. His importance is of course rarely reflected in the many historical overviews of philosophy produced since the middle of the 19<sup>th</sup> century, partly because he entered our collective memory as an epigon or systematizer of Leibniz. But even if this latter judgement were justified, it would still underrepresent the degree to which a systematization of the notoriously fragmented Leibnizian philosophy would have to be an original and philosophical effort of massive proportions. This reveals itself well in Wolff's rational psychology, i.e. his metaphysical theory of the soul.

In the *Vernünfftige Gedanken von Gott, der Welt und der Seele des Menschen, auch allen Dingen überhaupt*, also known as the *Deutsche Metaphysik*, first published in 1719, Wolff argued that the Soul is a simple substance rather than a composite, because no composite, and hence no body, could conceivably think (Wolff 1747: 463). For him, this realization is fruitful basis for many insights into the nature of the soul, because a great many things follow from being a simple substance, i.e. a simple independent thing:

A self-subsisting thing or a substance is that which has the source of its changes internal to it; a thing existing through something else, on the other hand, is nothing but a limitation of the former kind of thing. For example, our soul has a force through which it brings forth all of our thoughts one after the other in a steady order, and hence it is a self-subsisting thing. (Wolff 1747: 59-60; my translation)



It is noteworthy that Wolff introduces his notion of a self-subsisting thing or a substance with the example of the soul, even though he will only argue for the substantial nature of the soul many hundreds of pages later. This already suggests that the soul functions as the paradigm instance of a substance for him. Furthermore, he already suggests that the notion of a substance is intimately tied up with that of a force, understood as the internal principle of the changes of the substance: “The Source of the changes is called *Power* (Kraft): and therefore there is to be found in every self-subsisting thing a power of a kind that we do not encounter in other existing things.” (Wolff 1747: 60; my translation). For Wolff, it is mostly important that this force is an active principle, and not a merely passive one: “A power should not be conflated with a mere capacity (Vermögen); for the *capacity* is merely a possibility of doing something; since the power, in contrast, is a source of the changes, an effort (Bemühung) to do something must be found in it.” (Wolff 1747: 61; my translation). Hessbrüggen-Walter (2004: 58) traces this distinction between capacity and power to Leibniz, and concludes that the distinction rests on the fact that the former requires an external influence to be activated, whereas the latter contains its principle of activity (its effort) in itself.

In the fifth chapter of the *Deutsche Metaphysik*, Wolff argues that the specific power of the soul is a representative power (Wolff 1747: 469), more specifically the power to represent the world from the perspective of a specific material body (Wolff 1747: 489-490). From the fact that this power forms the essence or nature of a simple substance, he believes to be justified in concluding that there can be only one such representative power:

Since it is a simple thing, and there can be no parts in a simple thing, the soul cannot contain many distinct powers, because otherwise every force would require a specific self-subsisting thing existing for it alone. A power, after all, consists in an effort to do something, and therefore different powers would require different efforts. It cannot be, however, that a simple thing can have many simultaneous efforts, since it could no more move in different directions at once than a body that in its movement is to be regarded as an indivisible thing. And thus there is in the soul only one power, from which originate all of its changes, even though we are wont to call it by different names because of the different changes. (Wolff 1747: 463; my translation)

Wolff thus deduces from the simplicity of the soul that it can have only one power, and argues on this basis that there are no distinct faculties of the soul. There is something truly peculiar about Wolff’s analogy between a material body and the soul in the quoted passage, for he seems to say that a body that is under the influence of different forces would move in different directions at once. But this assumes that forces cannot

compose to form a single composite force. In fact, it was well known to natural philosophers in the 17<sup>th</sup> and 18<sup>th</sup> century that the parallelogram of forces constitutes a solution to how forces compose.<sup>2</sup>

The refusal to recognize multiple mental powers of course forces Wolff to explain how the one representative force can come to manifest itself in so different manners. The basis of his explanation is his contention that “all changes that can occur in a simple thing can be nothing else than variations in degree” (Wolff 1747: 56; my translation). The rationale behind this seems straightforward: since the force of a simple thing is unique, the changes cannot differ *qualitatively* (i.e. in the *kind of forces*), and since a simple thing is not extended, they cannot change *quantitatively* either. As a result, they can differ only in *degree of quality*. Indeed, Wolff goes on to argue, in the course of chapter 5 of the *Deutsche Metaphysik*, that all allegedly distinct mental faculties differ only in the degree of clarity and distinctness with which they represent, and that souls that seem to have different mental faculties actually differ only in the degree of clarity and distinctness of representation of which they are capable.

Wolff’s theory of the faculties thus comes down to the doctrine that every soul has just one faculty, because it is a simple substance, and that this faculty is rather a power, namely the principle of change, of the dynamics of mental life. This doctrine, however, ultimately feeds into a distinctly preformationist theory of knowledge and of mind:

Since the soul brings forth sensations through its specific power, the images and concepts of material things do not enter from the outside – on the contrary, the soul already has them in itself, namely in the manner in which it can as a finite thing have them in itself merely potentially, and not actually; and it merely develops them gradually in an order that corresponds with the body, from its own essence, by determining itself to actualize the potential. (Wolff 1747: 508; my translation)

Here, Wolff is admitting the link between his conception of the soul and the preformationist theory of knowledge Kant would go on to dismiss in the B-Deduction (cf. 3.3.2). This theory ascribes to the development of our knowledge an extreme determinism, since it not only assumes that all our representations are already contained in our soul as dispositions, but also that the development occurs according to the principle of sufficient reason. After all, Wolff constantly stresses that the succession of the degrees of clarity and distinction, the modulation of our mental life, is determined by this unyielding Leibnizian principle, and that force merely brings the

---

<sup>2</sup> Hessbrüggen-Walter (2004: 80) suggests that there is a methodological basis for Wolff’s claim, since it is easier to analyse the powers of the soul as manifestations of degrees of a single power rather than as the result of interactions between different powers.

determined to the fore. His constant demurals that the principle of determination is internal, did nothing to protect him from the charge of determinism that cost him a professorship.

### 6.2.2 Crusius and the Interanimation of Mental Powers

It is well-known among Kant scholars that the “All-destroyer” (as Moses Mendelssohn called him) was importantly influenced by Christian August Crusius, in all likelihood more so than by Christian Wolff himself. The originality and the specific influence of Crusius has often been overlooked because his philosophy is so intimately tied up with the Wolffian system, not because it is an extension of the latter (as Alexander Baumgarten’s seems to be), but because it is a *critique* of the latter. Because of this strong identification of the target, the terminology of Wolff is often adopted, even if only in order to transform and subvert it.

A strong similarity between Wolff and Crusius could, for instance, be wrongly inferred from the first pages of the latter’s *Pneumatologie*, the final section of his important and influential 1745 work *Entwurf der nothwendigen Vernunftwahrheiten wiefern sie den zufälligen entgegengesetzt werden*. The first chapter of the *Pneumatologie* is basically an extensive polemic against materialism in all its forms, much like Wolff’s rational psychology was. Like Wolff, Crusius denies that the specific functions of the soul, namely thought or representation, can be due to movement, and like Wolff he ascribes souls to animals, even though he is not willing to extend this courtesy to all living things, on the ground that it is a mistake to believe that everything that can suffer changes *perceives* as well. Because of this, he subscribes to preformationism against metamorphosis theories (1766: 891). Nevertheless, there is already a peculiar difference with Wolff, for although he agrees with the latter that thought does not consist in movement, he does believe the thinking subject to be *mobile* (1766: 876-877). An important difference is then that, for Wolff, the mind changes its perspective on the world solely by changing the proportion between clarity and obscurity and distinctness and confusion, whereas for Crusius this change of perspective reflects a real (and not phenomenal, however well-founded) spatial displacement.

But this difference is a humble one in comparison with the one announced at the beginning of chapter 2 of the *Pneumatologie*:

The understanding of a rational yet finite mind is not a sole fundamental force. One must instead picture it as a totality [Inbegriff] of certain fundamental forces and capacities, all of which share this feature, namely that they consist in a manner of thought and that they collaborate with consciousness to the advancement of knowledge of truth. I gladly admit, that we cannot possibly find out the precise number of the primary fundamental forces of our understanding.

This much is certain, however, that we would act in contradiction with the characteristics of fundamental forces if we were to regard the rationality of a finite being as a single power, which is merely singular in its nature. (1766: 907; my translation)

Crusius thus unambiguously rejects Wolff's theory of the fundamental representative force. The interesting aspect of this rebuttal is that it is done on the basis of the concept of a force, for Crusius claims that the identification of reason as a fundamental force would contradict the very concept of a fundamental force. We will thus have to look at Crusius' concept of a force insofar as it is pertinent to the present discussion.

In the more general sense of the word, a force is, according to Crusius, a possibility, connected to one thing, to bring about another thing (1766: 121). This general definition is already rich in positive and negative consequences. For one, Crusius says that a force already assumes causality and subsistence (1766: 121), since it is a capacity to causally bring something about, and that this capacity inheres in a subject. This allows him to say that no power or force is ever without a subject. But in characterizing force as a possibility, he is already diverging from the Wolffian orthodoxy, according to which a force is only that which is active, and the merely potentially effective is a capacity (*Vermögen*). For Crusius, these two terms only distinguish two stages of a force, and do not ontologically distinguish two kinds of properties:

As long as a force acts, and thus does that which is possible through it, and through which it is thought, it is called *living* [*lebendig*]. Of that, which was initially possible through the force, we say that it is henceforth present *actu secundo*, whereas it used to be merely present *actu primo*. Hence, to be present *actu secundo* means as much as really existing, insofar as we regard this existence as something which we credit to a force. To be present *actu primo*, on the other hand, means as much as being possible through an already really existing force, which nonetheless is not yet acting. (Crusius 1766: 124; my translation)

For Crusius, thus, force denotes both the mere potentiality and the reality, whereas for Wolff these two were radically distinct ontological kinds. An important consequence of this is that, according to Wolff, the representative force is determined to act only by its own striving, whereas for Crusius, as we will see, the activation (vivification) of a force can be dependent upon the agency of other forces.

But another major difference resides in the fact that Crusius enumerates a number of necessary conditions for a force to count as a fundamental force, in such a way that it leads directly to the falsity of the Wolffian account. Remember that, for Wolff, not just all the various kinds of activities and ideas, but all the individual activities and ideas of the soul are entirely due to the self-realization of the sole representative force. Crusius will charge this picture with incoherence, and will do so on the basis of his conception of a fundamental force.

Crusius needs the notion of a fundamental force in order to evade an obvious criticism, namely that his account is consistent with postulating for every possibility a separate capacity, which would come down to invoking occult qualities in the way that was mocked so efficiently by Molière in his *Le Malade Imaginaire*. Against this charge, Crusius admits that we should only allow for fundamental forces, in which other seemingly independent causal capacities are grounded (Crusius 1766: 128-129). This of course betrays a mechanistic commitment to an economy of causal principles. But he equally insists that such a fundamental force should be shown to be *really* capable of producing the necessary results, and in this he resists Wolff's tendency to simply assert the necessity of fundamental forces on metaphysical, epistemological or methodological grounds.

Crusius draws his readers' attention to 8 necessary characteristics of fundamental forces. In the first place, a fundamental force must *always* be present in the subject in which it inheres, and not just under specific circumstances (Crusius 1766: 131). This criterion reminds one of the Aristotelian distinction between essential and accidental properties. The second criterion is that, when we ascribe an effect to a fundamental force, we must be able to show how it causally follows from that force, and not be seduced to simply assume it for reasons other than the presence of a causal account (Crusius 1766: 131-132). It is not wildly conjectural to assume that, on this basis, Crusius would charge Wolff with simply postulating representative force as a simple fundamental force because of his metaphysical commitments, even though he can offer no plausible causal accounts of *how* representative force gives rise to such a variety of effects. The third criterion makes a similar suggestion, because it demands that a fundamental force have sufficiently uniform effects, i.e. that its effects can differ only in degree and direction. Of course, Wolff assumed that the degree of clarity and distinctness was all that was truly different between the various effects of the representative force, so he would not have been impressed by that complaint.

The fourth and fifth criteria are again better understood jointly. The fourth criterion postulates that the conditions for the activity or activation of a fundamental force should be found in the same subject as wherein the fundamental force itself inheres (Crusius 1766: 139). In other words, a fundamental force cannot be dependent, for its operativity, on conditions outside of its bearer. The fifth criterion states that the proximate effects of a fundamental force must inhere in the same subject as wherein the fundamental force resides (Crusius 1766: 140). Together, these two criteria state that there is a closure of proximate conditions and effects of a fundamental force under the subject of that force.

The sixth and seventh criteria specify that a fundamental force should not be derivable from another force, either causally (6<sup>th</sup> criterion), or logically (7<sup>th</sup> criterion) (Crusius 1766: 140-141). These two criteria are understandable, since they basically insist that the fundamental force be fundamental and not derivative. The 8<sup>th</sup> criterion is really

more an admonition on the basis of the previous ones, for it is a result of the maxim that there is no more in the effect than in the cause. Crusius puts this maxim to use in the 8<sup>th</sup> criterion by insisting that, where a force requires additional causes for its operation, and cannot itself produce these additional causes, we must conclude that there are also additional forces at work. Together with the second and third criteria, the eighth criterion states that anyone postulating a fundamental force must either give a plausible account of how this force can be responsible for all of its effects, or admit that there are other fundamental forces at work.

On the basis of these criteria Crusius mounts his attack against Wolff's theory of the fundamental force:

For [if this account is correct] we must either say, that the fundamental power to think, or to know the truth in general, and all that occurs in the understanding, would be merely determinations thereof [i.e. of this singular fundamental power]. The latter is perhaps true, but since the actions and changes of the understanding differ in more than merely direction and degree, this latter power is a mere general power, and not yet a true physical fundamental power. Or one should investigate whether from one of the parts listed above [i.e. one of the basic functions of the soul], which are required for reason, all the others can be derived through intelligible proximate causal effects, or whether they are at the very least so constituted that we cannot know whether they could not be derived from one if we were to have adequate knowledge. [...] Neither can be done, however. The power to continue an idea can perhaps be a determination of the action which constitutes the idea. It seems, as well, as if the capacity to abstract has its foundation merely in a higher degree of perfection of that activity through which the idea is thought in certain minds. I mean, it seems as if the capacity to analyze concepts, and to arrive from one idea to several of the many that are represented and contained in it, rests solely in a certain fineness and inner perfection of the original powers of thought by means of which the power to think brings forth this idea, which is possible through it, at once with a certain degree of inner perfection and the capacity to work in many ways. And when this is in fact the case, we cannot conclude to specific fundamental forces. Yet, we must keep in mind that every idea is an action, and that in every idea something different is represented than in any other idea; in other words that the proximate actions of a fundamental power would not continuously be similar, even though they must be, if all ideas are primarily dependent upon a single fundamental force. (Crusius 1766: 908-909; my translation)

Crusius is criticizing the recognizably Wolffian theory that all ideas in the mind are ultimately modifications of a single fundamental representative force on the grounds that such an account can make no sense of why the different accomplishments of this power are so diverse. He admits that ideas themselves may be modified in direction and degree, but he doubts that all ideas can be understood as directional and gradational

modifications of a single representation. Of course, he immediately warns that he does not thereby wish to conclude that the mind contains as many fundamental forces as it has ideas, for some reduction is possible. He merely concludes that, although we cannot know definitely how many fundamental forces the mind contains, we can be sure that it contains more than one on the ground that we are capable of thinking very different things (Crusius 1766: 909-910).

This allows for a very different picture of the dynamics of the mind, which I will call an interanimation of mental powers. As we have seen, Crusius believes the mind to contain multiple powers, and he believes powers or forces to be either merely potentially active, or really active. The question then becomes how the different powers of the mind become active throughout the course of our mental life, i.e. how they become “living forces”, how they become animated. Crusius suggests that the activation of one power is dependent upon the activity of another (Crusius 1766: 950-951), and that the subsequent modification of its direction and degree is also due to the interaction with another force (Crusius 1766: 947). In this way, the mind is not a constant self-developing of a single constant constitutive activity, as it is in the Wolffian picture, but rather a constant interaction of fundamental powers inhering in the same subject.

This final gloss is important, because it is supposed to allow Crusius to harmonize this interactive model with the fundamental nature of the forces and the non-composite nature of the soul. After all, Crusius did not demand of a fundamental power that it be completely unconditioned – he just required that its conditions lie in the same subject in which it also inheres. Given that all fundamental forces inhere in the soul, their interdependency does not therefore violate their status as fundamental. Furthermore, Crusius complains that he is not to be blamed if people attempt to infer from the multiplicity and interaction of the fundamental forces that the soul itself is composite. He states that such an inference would only be legitimate if the distinct and interacting forces were *moving* forces. Since they are not moving forces (per chapter 1 of the *Pneumatologie*), their “composition” is entirely dissimilar from the kind of composition that is typical for extended substances. In this way, he wants to bar the inference from the disunity of fundamental forces within a subject to the disunity of the subject.

With his account of the interanimation of mental powers, Crusius has offered a major alternative to the conception of the soul and its dynamics as understood by Wolff, and moved away from the monadological picture of the universe. This fits in his general project of putting the overly speculative Wolffian philosophy back on the ground by insisting on the distinction between the possible and the real. Crusius complained that the Wolffian school was mostly concerned with reasoning about possible objects, and consistently failed to show why these should be taken for real objects. He insisted that, whereas we are free to think up all kinds of objects, as long as they obey the laws of possibility, we cannot thereby conclude that they are also real. We can only encounter the real through *sensation*, for it is only in sensation that we are immediately forced to

take something as real (Crusius 1766: 28). This is because the capacity for sensation is a fundamental activity of the mind that cannot fail to be actualized when certain conditions are met. Sensation is thus the mark of the real because it occurs only when a certain capacity is realized, and this in turn can only occur when its conditions are realized. For Crusius, sensation is the mark of the real because it betrays causality, and all reality is grounded in causality.

### 6.2.3 Tetens on the Development of the Soul

In 1777, Johannes Nikolaus Tetens published his two-volume *Versuche über die menschliche Natur und ihre Entwicklung*. Although this work is dedicated to investigations into the nature of the soul, it differs from the works of Wolff and Crusius in being an attempt at *observational or empirical* psychology rather than *rational psychology*. Tetens aims to investigate the nature of the soul and its various functions insofar as they can be investigated through a purportedly scientific version of introspection. This general framework noticeably constrains the strength of the claims that can be made within it. As a result, Tetens is somewhat noncommittal on the issue of the unity or disunity of the faculties, however much it may pop up during his inquiries.

Tetens opens his first essay with a sketch of the issue of the unity of the faculties, commenting on the attempt to find a single fundamental force from which to derive all the many functions of the soul. In the preface to the book, he already announced his predilection for the approach of regarding the many different functions of the soul as quantitatively and intensively differing manifestations of the same fundamental force. Yet, he immediately warns that we cannot simply assume this picture, lest we overlook the many differences these manifestations might exhibit (Tetens 1777, I: xxvi-xxvii).

On the issue at hand, Tetens is both cavalier and reluctant, venturing wild hypotheses and then qualifying them with methodological remarks. For this reason, Dieter Henrich found that

Tetens gives all his divisions a provisional character. And if he ultimately attempts to find a hint of the basic power in the capacity of feeling common to all powers, then he himself has already taken that back in the introduction to his work, where he apologizes for every instance in which he gave in to the "propensity to a uniform system." Considering the unsystematic, merely descriptive character of Tetens's book, it is unlikely that Kant profited from it. (Henrich 1994: 25)

This judgement is too hasty, however, for Kant had plenty to learn from the wilder speculations of Tetens' work. The very title suggests that the work is not just, or not primarily, a piece of empirical psychology, but rather an investigation into the nature and development of the human being. This appears clearly from the final essay of the



first volume (the tenth essay), and even more so in the final, long division of the second volume.

In the tenth essay, Tetens again suggests, with great candour and little retraction (by his standards) the idea that, ultimately, the human soul must consist of a fundamental force (Tetens 1777, I: 730-731). But here, it soon becomes clear why he subscribes to such a theory, and why he still feels the need to revise it. For Tetens, it is imperative that we understand what it is that marks man off from the non-human, i.e. that we understand the nature of humanity. This means that we need to understand which capacities man has and that other (finite) living beings lack. Identifying such a capacity turns out to be difficult given the theory of the soul as simple and having only one fundamental force:

When the soul has to be understood as simple in the metaphysical sense, and as distinct from the organized body, then the question regarding the distinctive mark of the human soul leads us to two others. Is the development of the human soul a development of this incorporeal being, or does it consist solely in the development of its organized body, with which it is united? If it is not itself subject to development, to heightening or expansion of its faculties, then all its acquired dispositions consist merely in the brain's skill at serving the soul in its workings. What nature would the human soul then need? Indeed: none at all. The character of Man consists, on these conditions, merely in the special organization of the brain, or the representational machine. The soul of the waterpolyp, if one exists, as Mr. Unzer claims, will become a human soul. This has already been advanced by many respected new philosophers.

Or on the other hand: if in the inner nature of the soul developments and increases take place, then we can ask whether these also presuppose certain perfectable properties in its own special nature? If only the body would be concerned, and a dog's soul in a human body would develop humanly, without disposing of any other fundamental dispositions [Grundanlagen], than it had in the brain of a dog, then how could we inquire into the human soul at all? (Tetens 1777, I: 739; my translation)

Tetens' problem is thus that the traditional theory of the simplicity of the soul threatens to destroy the very idea of a specifically human soul. The idea of a simple soul suggests that a soul has its basic nature independently of its external circumstances, and that it has it as a simple nature. This again suggests that any soul has the same basic capacities as every other soul. The differences in the capacities of different souls are then completely due to the different bodies with which they are connected. Tetens sees this as resulting in a way from Bonnet's theory that the soul is the substantive entelechy of the brain, and that the soul itself has no specific nature. The problem with this is that there would be no specifically human soul, and more importantly, that the perfectibility of man does not reach his inner nature. Tetens believes man to be set apart from other creatures by his extremely perfectable nature, namely his capacity to develop himself,

overcome his earlier limitations and expand his capacities (Tetens 1777, I: 740). This idea he states to have derived from Rousseau (Tetens 1777, I: 742), but complains that it is hollow if we do not insist that it is the human soul, and not just the human body, that is subject to development and perfection, for otherwise none of these improvements would affect our inner being. Yet, do we therefore need to assume an original capacity for perfectibility in man?

Another problem is one that follows from the theory of the single fundamental force. This theory states that ultimately, the capacities of different souls differ only quantitatively, i.e. in degree, and not qualitatively. But this seems to imply, Tetens notes, that human souls have a higher degree of the same representative force. He suggests that this higher degree stems from the fact that the human soul has a higher degree of independent activity (*Selbsttätigkeit*), and that is therefore more plastic (*modificierbar*). The specific difference of the human soul is then its high degree of independent activity. This idea seems to be a gloss on the Wolffian idea that the fundamental principle of the soul is an activity rather than a potency. Tetens now seems to say that, much like the capacity for representing clearly and distinctly depends on the degree of the fundamental power of representation, the capacity to reach higher degrees of power of representation depends on the intensity of the *conatus* driving the fundamental force.<sup>3</sup>

In the second division of the second volume of the *Versuche über die menschliche Natur und ihre Entwicklung*, Tetens presents a more lengthy and systematic account of his own theory of the internal development of the human soul. He starts off by arguing, on the basis of what he calls observational data, but what we would now consider anecdotal material, that the growth and development of our faculties is more than the mere increase of the stock or the series of ideas of which we dispose. Against this more narrow conception of development, he advances that specific activities of reflection and representation contribute to the development of the faculties themselves. For the active faculties of the mind, specifically the understanding, this means that we acquire a greater capacity to conjure ideas of objects, and to conjure ideas of our representational activities. The latter increase also allows us to recall past representations more easily. (Tetens 1777, II: 390-391). For the passive or receptive capacities, such as sensibility, the increase consists in a greater receptivity to the nuances of nature in sensation, as well as a greater capacity to be modified, to store, past sensations. (Tetens 1777, II: 415-416).

---

<sup>3</sup> If this reading is correct, it seems that Tetens needs to notionally distinguish the *conatus* of a force from the force itself. It is unclear to me whether one can do this on Wolffian grounds, i.e. distinguish the intensity of the force from the intensity of its striving. It seems artificial to me, since it seems to require that a power is a bare capacity after all, which requires an external factor, its *conatus*, to be active, whereas Wolff seems to insist that, in the case of a true force, the *conatus* is not distinguishable from the force itself.

Finally, Tetens insists that an increase in the capacity of one faculty ultimately also causes the increase of other capacities.<sup>4</sup>

This whole theory of a genuine development of the soul raises an important question, namely how we should understand development here. It is in this context that Tetens discusses the distinction between epigenesis and preformation in order to find which theory understands development better, and whether there might be a difference between the process governing the development of the soul and the process governing the development of the body.

Tetens initially offers this general hypothesis concerning the development of the soul, which he assumes to be neutral with regard to the issue of preformation or epigenesis:

This is then the law of the development of man in his psychic nature, and it is similar to that, according to which the body grows. Feelings and representations are the nutritional juice [Nahrungssaft] that are supplied to the fundamental power, which is initially limited in its activity to Feeling, and , thereby stimulating and activating it. Every active expression of this force strengthens it. That which was only predisposition [anlage] or possibility, becomes disposition [Disposition], capacity [Fähigkeit], drive [Trieb] and Skill [Fertigkeit]. The transition from mere skill to the nearest predisposition or disposition rests, according to the common opinion, on the fact that something additional is required to arrive from the former to the latter. (Tetens 1777, II: 426-427; my translation)<sup>5</sup>

In this picture, the soul initially has mere possibilities, which are then activated through the intake of experiential data. This raises several issues, of course. First of all, is the effect of the experiential intake limited to occasioning the activation, or does it influence the mind in other ways as well? Secondly, are we to understand the process in the preformationist way and regard the faculties emerging later in the process of development as already preformed in the initial stage, or are we to regard them as new structures emerging through the process of development? Tetens himself raises these questions in the following fascinating passage:

When any increase of the soul on one side radiates across the whole of the soul, and activates its power in other directions as well, then what does it effect? Does

---

<sup>4</sup> This is of course understandable from both the Wolffian and the Crusian perspective. For Wolff, the increase in one faculty would imply an increase in all others because they are ultimately the same faculty. For Crusius, the increase in one faculty would imply an increase in others because of the interanimation of the faculties.

<sup>5</sup> I believe the last sentence is a confusion on Tetens' part, because the "Fertigkeit" is intended to be a highly active version of an "Anlage" or "Disposition". I therefore suggest we read the passage as saying that something external is required to pass from the "Anlage" to the higher degrees of "Thätigkeit".

it produce a new faculty, or does it merely strengthen and maintain the faculty that is already present in nature in such a way, that it can express itself and go forth as soon as a favorable situation comes along? That is in other words the question that has been investigated with exceptional zeal as far as the body is concerned, and that has not even been raised as far as the development of the soul is concerned: namely, whether the articulation of the faculties is merely an *evolution* of preexisting natural predispositions [Naturanlagen], or an *epigenesis*, that brings forth new faculties, of which nothing more preexisted than the receptivity to acquire them. The German philosophers are almost invariably epigenesists of the soul, much like the German physiologists are evolutionists [i.e. preformationists] of the body. Hutcheson, Reid, Beattie, Oswald, mostly however Home, attribute many innate fundamental feelings to man. [...] We can regard those that consider such determinate feelings as defenders of physiological evolutionism, since for them, the predispositions to the different manners of sensitivity or activity of nature, must be originally present in miniature next to one another in the soul, in the same way that in Bonnet's opinion, in the fertilized egg and in germinating seeds all the canals and vessels of the whole body must be present in both form and predisposition. (Tetens 1777, II: 434-435; my translation)

This passage is fascinating first of all because it starts off with the claim that the previous picture of development, and the very term *entwicklung*, is neutral with respect to epigenesis and preformation. This means that Tetens regarded the initial account as not already deciding whether epigenesis and preformation is to be preferred. Secondly, it offers a peculiar taxonomy of positions on the nature of the soul, since it regards the German theories of the soul as more epigenesist than the empiricist theories of the British and Scottish authors. This requires some comment, because it seems as if Tetens does so on the basis that German philosophers attribute to the soul higher order capacities that develop later on and have generic functions, whereas empiricists posit a special feeling or capacity for every higher-order capacity. Physiological epigenesis would then be the theory that a single principle or a limited set of principles suffices to generate the various functions of the soul, whereas physiological preformationism would require that each function of the soul stems from a distinct preexisting origin.

As usual, Tetens seeks to find a middle ground between the two theories, in the form of a compromise. He suggests that the physiological preformationists have a good point in insisting that not all capacities can emerge from a single source, given the fact that the capacity for reasoning, for instance, differs infinitely from the capacity for feeling (Tetens 1777, II: 445). For this reason, he allows that there might be multiple predisposed capacities, but insists that we nevertheless also attempt to explain non-fundamental capacities as newly emerging during development. The resulting theory is that the development of some preformed elements can give rise to novelty, or, as Tetens expresses it, that there can be epigenesis through evolution.

Tetens thus believes that the development of the soul proceeds according to a process analogous to the development of the body. He therefore first tries to develop a theory of generation of the body, not on the basis of research or observations, but on the basis of a review of the literature. He distinguishes three global positions: the system of evolution, which we know as preformation, defended most recently and most capably by Haller and Bonnet, the position of epigenesis, which he associates with Caspar Friedrich Wolff, and finally a group of related positions that are neither, defended by Buffon and Needham (and, as he adds later, Maupertuis) (Tetens 1777, II: 448).<sup>6</sup> He also importantly notes that the first two positions agree on two principles. The first principle is basically the version of the principle of sufficient reason Leibniz called the principle of causality, namely that nothing exists without ground or cause. In the context of embryology, this means: “the germ [Keim] and its internal and external properties, and that which is added to it, jointly contain the full sufficient ground of development, and determines the internal form, size and construction of the body that is produced” (Tetens 1777, II: 453; my translation). In the first two chapters of this dissertation, I already indicated how the issue of the principle of sufficient reason interrelated with theory of development in early modern philosophy, and in Kant’s philosophy prior to the 1770s. Here, Tetens uses it to bracket or even dismiss the third position he distinguished, because it does not sufficiently explain the grounds (Buffon), explains without sufficient ground (Needham), or introduces sufficient ground through quasi-intelligences (Maupertuis) (Tetens 1777, II: 464-473). This leaves him with two viable alternatives: evolution and epigenesis, in the versions of Bonnet and Wolff respectively.

Tetens also discerns another principle on which both preformationism and epigenesis agree: the principle that an organized being develops from a germ (Tetens 1777, II: 455). This would seem puzzling to those readers who take the notion of a germ to be “unambiguously preformationist”, since Tetens considers the term neutral with respect to theories of development. It should not be so surprising, however, since it resonates with my discussion of chapters 2 and 3, where I argued that Kant also used the terms germs and dispositions in a theory-neutral manner, and saw them as nothing more than the constraints on the process of development such that the latter can follow a determinate pattern or lead to a determinate result. Similarly, Tetens takes the term *Keim* to mean any causal influence on the process of development that accounts for its determinacy or specificity (Tetens 1777, II: 475-476). The difference between preformation and epigenesis then amounts to the fact that the former insists on the

---

<sup>6</sup> Remember that I argued, in chapter 2, that this was not an exceptional position at the time. Kant made very much the same distinction even before Tetens published the work I am discussing here.

uniqueness of germs, namely the original germ in which all forms are inchoately present, whereas the latter allows for multiple germs, some of which are external, and for the emergence of new forms through development.

Tetens' presentation is obviously meant to make the two leading rival theories somehow commensurable, and to allow him to propose a synthesis. This synthesis is the theory that a preformationist development *can* produce new forms, contrary to what authors like Bonnet insist (Tetens 1777, II: 497). He argues for this by starting off from Bonnet's theory and then showing that this theory does not preclude the production of new forms. Bonnet's theory holds that there is no new production of forms in development, that there is only growth and expansion of existing forms. This growth and expansion occurs according to intussusception – as Tetens remarks, both preformationists and epigenesists agree that the nutrition takes the form of intussusception into original material (Tetens 1777, II: 495). However, in taking in new materials and in growing, preexisting forms can not only be significantly altered: new connections between existing forms become possible, namely connections between the newly integretated parts. If such connections can themselves constitute forms – and Tetens believes they can – then new forms can arise simply through the preformationist process of development, i.e. a preformationist development can have epigenetic effects.

This theory allows Tetens to distinguish between forms that were present in advance, namely those for which there were germs present, and forms that arise only later, through the epigenetic process, and did not arise out of preexisting germs. The former are essential, the latter accidental forms of the organization. Nevertheless, this accidental nature is not entirely the effect of mere outside influence: there are in the germs some tendencies, however small, to prefer one new formation over another (Tetens 1777, II: 521-522). Tetens constantly argues that even epigenesis must assume some such constraints to be in place on the process of development, and that it too requires that new development proceeds from some previous structure, even though the previous structure may be relatively dissimilar. He also argues that, whenever Bonnet resists any straightforward caricature of preformation, he is bound to admit great transformations in the process of development. In doing so, he of course undermines his claim that these transformations are merely increase and growth.<sup>7</sup>

---

<sup>7</sup> Tetens also insists that *new germs* can be produced in nature. He does so by using Buffon's theory of the origin of reproductive materials. Buffon hypothesized that the reproductive materials present in the reproductive organs are produced by the organs, and that they form the inchoate forms of the same organs in the offspring. Tetens notices that this theory is perhaps compatible with preformation, because it states that while new germs are produced, they are produced by the things of which they are the germs in the parent organism.

Having provided this general picture of his theory of “epigenesis through evolution”, Tetens can then apply it, through analogy, to the development of the soul. Of course, he insists that analogies have their limits, but he also indicates that the existence of the analogy might explain why the soul and the body, despite their distinct natures and their separate developments, follow the same pattern of development. Nonetheless, this analogy goes very far: Tetens suggests that the human soul can itself, though it is immaterial, be produced by the act of fertilization, and he insists that it is already in its embryonal stage distinctly human. Its distinct nature consists in the fact that it already has the germs for certain properties, and that it has more or less tendency to be modified in certain ways. In this context, Tetens first notices that this is his solution to the debate on innate ideas. He argues that, although he wants to believe that we in fact do receive information from the outside, and are not born with all ideas innately present, we must be born with the disposition to receive certain impressions and not others, to be sensitive in some ways and not in others (Tetens 1777, II: 545).

Tetens thus argues that the soul develops through an epigenesis prompted by evolution. In the case of the body, this was done through the intussusception of nutrients into the existing organic forms. In the soul, it is through the intake of sensations into the powers of the soul. (Tetens 1777, II: 548) This explains why the development of the soul does not solely consist in the increase in the stock of ideas, since the latter increase can strengthen the powers of the soul and give occasion to new connections. Finally, it allows Tetens to refine his theory of the distinction between fundamental and derivative faculties: the fundamental faculties are those that, like the fundamental forms in the body, are present in advance as a germ, whereas the derivative ones are those that are due to the growth and strengthening of the faculties, or composites allowed for by the new connections that arise between existing faculties as these expand.

Given all this, it is wrong to suppose, with Henrich, that Kant would not have learned much from Tetens, for although the latter did not dare to make hard claims about the debate on the unity of the faculties, he did insist that it connects with the idea of the development of the mind. To be sure, Wolff and Crusius also had their pictures of the dynamics of mind: for Wolff, it consisted in the continuous effort of the unique representative force developing its internal determinations over time, and for Crusius it consisted in the interanimation of powers in the soul and their mutual occasioning. Tetens, however, made the connection with theories of development explicit, and also revealed the merits of an epigenesist picture of the mind. It is clear that Kant was fascinated by the latter, and that he too wished to have a developmental conception of the faculties. It would be wrong, however, to assume that he simply adopted Tetens theory, since the very shift from the context of rational(ist) psychology to that of transcendental epistemology brings along important changes. Moreover, there are subtle but nevertheless important differences between the Tetensian and the Kantian

accounts of epigenesis. In the final chapter of this dissertation, I will discuss Kant's conception of the epigenesis of the faculties. In the remainder of this chapter, I will present my interpretation of his conception of the system of faculties.

## 6.3 The Disunity of the Gemüt

### 6.3.1 The Distinctness of Sensibility and Understanding

As we have seen, the distinctness of sensibility and understanding is a core doctrine of Kant's critical philosophy. This suggests that, for Kant, the two faculties do not ultimately arise from the same source, and are not different manifestations of the same underlying capacity. In this subsection, I will discuss Kant's reasons for maintaining that the lower and the higher faculties are really distinct, reasons that are ultimately of an epistemological nature as well.

The obvious target of Kant's critical insistence on the distinction between sensibility and understanding is the Leibnizian-Wolffian theory of ideas. According to that theory, the ideas of perception and those of conception differ only in the degree, and not in kind: perceptions are often confused to a high degree, because in perceiving particulars we have a full idea of them, but do not discern all the various ideas contained in this full idea. In the previous chapter, I already mentioned that, according to Leibniz, a full idea of a particular does not just consist of an enormous amount of different ideas, it consists of an *infinite amount* of different ideas. This is due to the fact that, in Leibnizianism, identity, and therefore individuality, is governed by the principle of the identity of indiscernibles. According to Leibniz, two things that have exactly the same properties, conceptually, are identical, so that, ideally, one could grasp an individual through a concept, namely its full concept, the concept that characterizes it, and only it, precisely.

Kant denies that it is possible to grasp individuals through mere concepts. According to him, we are only ever presented with individuals in sensibility, and not in the understanding. This means that sensibility does play a notionally distinct role in thought, contrary to what McDowell seems to maintain. However, this in turn suggests that Kant might be subject to the Myth of the Given, and that therefore the theory of the distinctness of the faculties might be intended, but nonetheless fundamentally misguided. I will argue that this latter conclusion can be avoided if we realize that it is its *intuitive* nature, and not its *receptive* nature, that is responsible for sensibility's distinct contribution to knowledge. First, however, we must dismiss some initially plausible, but ultimately unsatisfying, alternative interpretations.



The first interpretation assimilates Kant's reasons for maintaining the distinctness of the faculties to those of Crusius. In subsection 6.2.2, I mentioned that Crusius criticized Wolff for not observing the distinction between the possible and the real in philosophy. Concretely, he claimed that whenever Wolff spoke of entities, objects and things, he was actually dealing with mere concepts of possible things. In order to be able to pass from the concept of a possible thing to the thing itself, something further is required, and this further element is *sensation*. For Crusius, the real always refers to a causal ground that is responsible for activating the merely potential. In absence of the causal influences of other objects, i.e. the sensations they cause in us, we are not entitled to conclude that our concepts or representations have reality.

If Kant were to hold this theory, it would mean that he believed the foundation of our thought about the real to be constituted by a causal relation with the external world that provides us with sensory particulars, in *casu* sensations. This, however, is the very paradigm of the Myth of the Given, so we should hesitate to ascribe it to Kant directly. Moreover, as I have already indicated in section 6.1, the attribution to Kant of any serious claim about the causal ground of sensations has two demerits: first, its textual basis is only a possibly provisional phrasing to be found at the very beginning of the *Transcendental Aesthetic* that is never theoretically articulated; second, it leads to a highly intractable dilemma within transcendental idealism. There are thus good reasons to look for other interpretations of the sensibility-understanding distinction.

The second interpretation holds that it is not the causal *origin*, but rather the specific *content* of perceptual information that yields its non-conceptual nature. A traditional way of expressing this is that sensory content has a fineness of grain not to be had through conceptual content, that, in Fred Dretske's way of putting it, sensory content is *analog* whereas conceptual content is *digital* (Dretske 1981: 141-142). This means that sensory information takes the form of continuous quantities, whereas concepts make discontinuous cuts in these continuous degrees, and thereby "abstract" from the particular, "classify" where before there was only a difference of degree. I believe that some such reasoning is present in Kant, but that he did not believe sensory information to simply *be* continuous. In fact, I take the argument of the anticipations of sense-perception to imply that the categorical framework requires that sensation takes the form of the continuous. In absence of the category of quality, there would be no reason to expect the sensory to be analog. I will not here go into the difficult argument for this, and simply indicate it. In any case, although I believe that Kant can hold that sensibility and understanding differ like the analog and the digital without falling prey to the Myth

of the Given, I do not believe it is because it is *analog* that sensory information is information of particulars.<sup>8</sup>

A third interpretation starts from the familiar Kantian idea that the real needs to be the object of possible experience, and interprets the idea of possible experience as the idea of possible perceptual acquaintance. This theory is an advance on the former because it does not actually require that something is perceptually given; it only requires that we can possibly encounter something. This vastly expands what we can count as “real” beyond the *petites histoires* of our own perceptual episodes. This theory amounts to what Gareth Evans has criticized as Ideal Verificationism, which he thinks amounts to the following idea from Berkeley’s *The Principles of Human Knowledge* (2<sup>nd</sup> edition 1734):

For the question, whether the earth moves or no, amounts in reality to no more than this, to wit: whether we have reason to conclude from what has been observed by astronomers, that if we were placed in such and such circumstances, and such and such a position and distance, both from the earth and sun, we should perceive the former to move among the choir of the planets, and appearing in all respects like one of them; and this, by the established rules of nature, which we have no reason to mistrust, is reasonably collected from the phenomena. (Berkeley 2009: 107)

Evans suggests that this idea is an illegitimate interpretation of perfectly legitimate philosophical clues:

Who can deny, with his hand on his heart, that such imaginings play a part in both his ordinary and his philosophical thinking? But it is not the imaginings that need to be denied; only the interpretation placed upon them. Evidently what is essential for a subject's conception to involve a spatio-temporal object is his conceiving that *somewhere* there exists an object which his thought concerns, and these imaginings are no more than the reflection of this idea. Our thought about the spatial world is, perhaps necessarily, accompanied by models or maps. Most of us have, for example, such a model of the solar system, with the sun in the centre, and the planets revolving at differing distances around it. The mistake which renders some form of ideal verificationism almost inescapable-is to suppose that,

---

<sup>8</sup> Mind that the analog-digital distinction might be harmonized with Leibniz by stating that the information in sensory perception could be grasped conceptually only through an infinitely fine-grained concept. I suspect that this is part of the reason why Kant requires that quality be conceived as analog in nature, namely because it could then be grasped conceptually in principle, through an infinite process of refinement and conceptual elaboration. Regarding quality through continuous quantity makes quality conceptually graspable without reducing it completely.

in constructing these models, we are thinking of the *content of a possible experience*.  
(Evans 1982: 99-100)

Instead, Evans suggests that what these clues point to “is not ideal verificationism, but the crucial role of our conception of the spatial world in much of our thinking” (Evans 1982: 100).

It would initially seem as if Kant’s usage of the idea of possible experience bears witness of his falling prey to the misconception against which Evans warns. This conclusion, however, is precipitous, for we may perhaps understand Kant better by taking our cue from Evans himself, and understand the concept of “possible experience” not through “possible perceptual acquaintance”, but through the concept of spatiality involved. In fact, as I will now argue, Kant may have believed our capacity for demonstrative identification, i.e. of non-conceptual identification of – and hence reference to – particulars, to depend not on these particulars being possible content of our experience, but on the *a priori* forms of sensibility, namely orientable space-time.

The distinction between the sensible and the intelligible was a main topic of Kant’s thought throughout the 1760s. These inquiries culminated in the 1770 Inaugural Dissertation, which is explicitly on the topic – as is revealed by its official full title *De Mundi Sensibilis atque Intelligibilis Forma et Principiis*. This does not mean that the distinction would pass on unchanged to the critical writings, since it runs together those between concept and the intuition, between the real and the logical, and between phenomenon and the noumenon. Tracing the various transformations which have led to the differentiation of these three oppositions characteristic of the critical Kant, is a task beyond this dissertation. Here, I will merely indicate one major line of thought that runs through the discussions from the 1760s and that sheds light on the distinction between sensibility and understanding from the 1780s onwards.

In 5.3.1, I have already mentioned in passing one argument for the distinction. In the 1763 *Only Possible Argument* text, Kant briefly invoked the distinction between logical and real opposition in order to distinguish his own position on the *ens realissimum* from that of Leibniz. In a paper of the same year, *Attempt to Introduce the Concept of Negative Magnitudes into Philosophy*, Kant further articulates that distinction and discusses the various philosophical uses to which it can be put. Remember that, for Leibniz, two concepts are either compatible or contradictory. They are contradictory if one predicates something of a subject which the other denies of it. This opposition is binary, in the sense that one is affirmative, and the other negative, and that there is no other option (the law of excluded middle). If two contradictory predicates are predicated of the same subject, then that subject cannot be thought (is *incogitabile*) and cannot exist. (AA II: 171)

Kant argues that, besides this logical opposition holding between contradictory predicates, there is another opposition, namely real opposition. Really opposed

predicates, unlike logically opposed ones, *can* exist simultaneously in the subject without abolishing the very possibility of that subject. They do, however, at least partially cancel each other out. (AA II: 171-172) Kant makes this understandable through the concept of negative magnitudes in mathematics (AA II: 172ff). Negative numbers are something positive in mathematics: they have a determinate magnitude, and are therefore distinct from zero. Yet, they are opposed to positive quantities in that positive and negative quantities cancel each other out. For instance, the sum of +5 and -5 is zero, and not a contradiction.

Kant's claim is that the concept of a negative quantity provides a different kind of opposition. Whereas in the case of logical opposition, there is typically one truly positive, and one merely privative member of the pair, and the two together lead to an impossibility, in the case of real opposition, both members are truly positive, and neither merely privative, and they do not jointly cancel out the possibility of the subject in which they inhere. The concept of a negative magnitude makes the latter trait easy to understand, but we also need to understand why the former trait, that of being equally positive, follows from it. In what sense can negative numbers be considered equally positive as positive ones?

We can understand this better if we take a look at one of the examples Kant gives, namely that of opposed forces. If a given material object is subject to two opposed forces at the same time, this does not involve a contradiction. What rather happens is that the two opposed forces cancel each other or that the weaker one cancels out a magnitude of the stronger equal to its own magnitude. In the former case, the object remains at rest, even though it is subject to forces – rest does not imply absence of force. In this example, we can not only see what Kant means by non-contradictory opposition, we can also see why he considers both opposed members of the pair to be equally positive: they differ in sense or direction.

This reveals further why Kant thinks of negative numbers as positive, namely by associating the minus-sign not with privation, but with a difference in sense. Along the single dimension of the series of numbers, + means away from the zero-point in one sense, whereas – means away from the zero point in the opposed sense. Hence, adding a negative magnitude to a positive one may be considered as the counteracting of the positive vector by the negative vector. According to Kant, arithmetical addition is what we would now call vectorial addition, but in a one-dimensional system (along a single axis). Hence, real opposition is associated with a vectorial, and therefore directional, nature, and even arithmetics is best understood through the model of forces.<sup>9</sup>

---

<sup>9</sup> I believe we can understand, on the basis of this conception of arithmetic, how Kant came to the idea that, whereas geometry derives its formal content from the a priori intuition of space, arithmetic and algebra

The concept of real opposition resonates with the central point of another text from the 1760s, the 1768 *Concerning the ultimate Ground of the Differentiation of Directions in Space*. In that text, Kant seeks out the condition of the possibility of comprehending the very idea of a direction, and concludes that it is the concept of absolute space. The argument of the paper seems to be directed at Leibniz, who believed that space is a mere *phaenomenon bene fundatum*, and that spatial relations can be reduced to the non-spatial properties of things. Kant attacks this assumption with the so-called incongruent counterparts argument. The argument of incongruent counterparts suggests that there may be two objects which are perfectly similar as far as their conceptual description is concerned, but that can be different as far as their specifically spatial properties are concerned. Kant's example is that of the left and right hand, but is better understood as that of two objects that are each other's perfect mirror image. According to Leibniz, two objects which are conceptually indistinguishable are identical (are not really two). Kant argues here that there are non-conceptual, i.e. irreducibly spatial properties in virtue of which an object can be distinct from another whilst being conceptually indiscernible from it, and these properties are the orientable, directional properties objects have by being situated in absolute space.

---

derive their formal content from the a priori intuition of time. The idea can be better understood through a later elaboration of Kant's philosophy of mathematics by William Hamilton. Hamilton presented his own version of the Kantian theory in a preliminary essay to one of his mathematical papers published in the transactions of the royal Irish Academy. Like Kant, he finds fault with the usual way of thinking about the contrast between positives and negatives, and introduces his theory in order to remove "the difficulties of the usual theory of Negative and Imaginary Quantities, or rather substituting a new Theory of Contrapositives and Couples, which he considers free from those old difficulties, and which is deduced from the Intuition or Original Mental Form of Time: the opposition of the (so-called) Negatives and Positives being referred by him, not to the opposition of the operations of increasing and diminishing a magnitude, but to the simpler and more extensive contrast between the relations of Before and After, or between the directions of Forward and Backward; and Pairs of Moments being used to suggest a Theory of Conjugate Functions, which gives reality and meaning to conceptions that were before Imaginary, Impossible, or Contradictory, because Mathematicians had derived them from that bounded notion of Magnitude, instead of the original and comprehensive thought of Order in Progression." (Hamilton 1837: 297) The idea behind this remark is clear: as long as we think of numbers as magnitudes, we can make little sense of negatives, for negative magnitudes do not seem to make sense. If we instead think of numbers as the members of an ordered series, we can allow for different places and movements within that series. The fundamental concept of mathematics is therefore order, and not magnitude. Hamilton goes on to suggest that the capacity to think an ordered series depends on the concepts of order in space and time: "The notion or intuition of Order in Time is not less but more deep-seated in the human mind, than the notion of intuition of Order in Space; and a mathematical Science may be founded on the former, as pure and as demonstrative as the science founded on the latter" (Hamilton 1837: 297). I believe this was Kant's intuition as well: the very idea of an ordered series depends on an irreducible grasp of the concept of order, and this order lies in the orientability of time, as, for instance, the very capacity to count relies on the capacity to orient each step of counting in a synoptic whole of time.

The 1768 argument falls of course short of the critical theory, because, although it admits that the concept of space is indispensable to our capacity to orient ourselves and to conceive ourselves as related to other things otherwise than by logically conceiving of them, it assumes that space is a *concept*, and not an intuition. Secondly, Kant is still in doubt over the contribution of sensibility to our thought. This passage from the *Inaugural Dissertation* reveals an advance:

There is (for man) no *intuition* of what belongs to the understanding, but only a *symbolic cognition*; and thinking is only possible for us by means of a singular concept in the concrete. For all our intuition is bound to a certain principle of form, and it is only under this form that anything can be *apprehended* by the mind immediately or as *singular*, and not merely conceived discursively by means of general concepts. But this formal principle of our intuition (space and time) is the condition under which something can be the object of our senses. Accordingly, this formal principle, as the condition of sensitive cognition, is not a means to intellectual cognition. Moreover, since it is only through the senses that all the matter of our cognition is given, the noumenon as such cannot be conceived by means of representations drawn from sensations. Thus the concept of the intelligible as such is devoid of all that is *given* in human intuition. The *intuition*, namely, of our mind is always *passive*. It is, accordingly, only possible in so far as it is possible for something to affect our sense. Divine intuition, however, which is the principle of objects, and not something governed by a principle, since it is independent, is an archetype and for that reason perfectly intellectual. (AA II: 386-397)

Kant's theory here moves a considerable distance towards his position in the *Critique of Pure Reason*. First of all, he notes that our form of knowledge is essentially divided between sensibility and understanding, between the intuition and the concept. He also already contrasts the *intellectus ectypus* and the *intellectus archetypus*, although here he may be more committed to the position that there not only *is* such an intuitive intellect, but that it is also the source of the existence of the thing in itself. This standpoint still appears in the famous letter to Marcus Herz, although there it already reveals some of the evolutions of Kant's thought on the subject.

The most important feature of the quoted passage is, however, that it distinguishes between two aspects of sensibility, namely that, for all its passive nature, sensibility makes a contribution of its own, namely the form(s) of intuition, space and time. Moreover, Kant suggests that it is through this form that we are capable of grasping things as *singular*, rather than as general. It thus seems that sensibility has two features: its receptivity, which makes us capable of acquiring the matter of thought, and its form, which allows us to grasp and think particulars.

I believe this sets us on the path of understanding the critical account of sensibility. Kant defines sensibility as a faculty that is receptive and intuitive. It is receptive

because, through it, we are *given* information, and it is intuitive because it presents information in terms of particulars. The contribution of sensibility can then be understood as allowing us to have perceptual judgments or beliefs as judgments or beliefs *about particulars*. Traditionally, the distinct role of sensibility has been taken to be its receptivity, but we have seen that this is problematic. I suggest instead that it is its *intuitive* nature that makes up its contribution to thought. In itself, it might be that sensations could be directly given to the understanding as its matter. But this is not how Kant pictures it: sensations are only presented to the understanding through the form of intuition: space and time. I therefore suggest that the specific contribution of sensibility is its intuitive nature, that is, its capacity to have thought relate to particulars, and that in turn this capacity depends on the intrinsically and irreducibly orienting and orientable nature of intuitive space-time.

This makes sense of various aspects of Kant's presentation of the theory. First of all, Kant is unerringly silent about the receptivity of sensibility, and focuses instead on a proof for the intuitive nature of space and time. He thus seems more concerned with showing that the contributions of space and time are not reducible to conceptual capacities. It also reveals why Kant insists that in absence of intuitions, we are only ever presented with concepts, and never with objects. In the "Amphiboly of Concepts of Reflection" from the Critique of Pure Reason, he again lists the vectorial nature of real opposition and the orientability of space as reasons why Leibniz is wrong to believe that we can relate to objects merely by having the concepts of those objects:

The principle of indiscernibles is really based on the presupposition that if a certain distinction is not to be found in the concept of a thing in general, then it is also not to be found in the things themselves; consequently all things are completely identical (*numero eadem*) that are not already distinguished from each other in their concepts (as to quality or quantity). But since in the mere concept of anything abstraction is made from many necessary conditions of an intuition, it is with peculiar haste that that from which abstraction has been made is taken as something that is not to be encountered at all, and nothing conceded to the thing except what is contained in its concept.

The concept of a cubic foot of space, wherever and however often I think it, is in itself always completely the same. Yet two cubic feet are nevertheless distinguished in space merely through their locations (*numero diversa*); these are conditions of the intuition in which the *object* of this concept is given, which do not belong to the concept but to the entire sensibility. In the same way, there is no contradiction at all in the concept of a thing if nothing negative is connected with something affirmative, and merely affirmative concepts cannot, in combination, effect any cancellation. Yet in the sensible intuition in which reality (e.g., motion) is given, there are conditions (opposed directions), from which one had abstracted in the concept of motion in general, that make possible a conflict, which is certainly not a logical one, that produces a zero = 0 out of that which is entirely

positive; and one could not say that all reality is in agreement just because no conflict is to be found among its concepts. According to mere concepts the inner is the substratum of all relation or outer determinations. If, therefore, I abstract from all conditions of intuition, and restrict myself solely to the concept of a thing in general, then I can abstract from every outer relation, and yet there must remain a concept of it, that signifies no relation but merely inner determinations. (CPR A 281-283 / B 337-339)

Leibniz's doctrine that we can relate to objects by concepts alone is false because what we would be related to are not the objects, but rather the general concepts through which we think them. In order to have an object as a particular before us, and not just a concept, we need the forms of sensibility, and their specific relational and directional features.

The suggested reading also sheds light on the role accorded to imagination in the Transcendental Deduction: there, imagination is understood as *intuitive yet spontaneous*. This can be understood as follows: understanding is not of itself related to objects, and sensibility cannot conceive at all. Thus, in order to bring the two together, a faculty is required that allows us to *think about particulars*, having the "think" aspect due its spontaneity, and the "about particulars" aspect due to being intuitiveness. It would thus be wrong to see the imagination as the shared origin of sensibility and understanding: it is rather a mediating capacity that is necessary because of the *chasm* between the two capacities:

We therefore have a pure imagination, as a fundamental faculty of the human soul, that grounds all cognition a priori. By its means we bring into combination the manifold of intuition on the one side and the condition of the necessary unity of apperception on the other. Both extremes, namely sensibility and understanding, must necessarily be connected by means of this transcendental function of the imagination, since otherwise the former would to be sure yield appearances but no objects of an empirical cognition, hence there would be no experience. (CPR A 124).

As I understand this passage, an object requires two aspects: a particular which is characterized, and a concept which characterizes it. Jointly, they present us with beliefs or judgments of the kind: "This is an  $x$  which is  $F$ ". Here,  $x$  stands for the base concept of a concept in general, characterized by the categories. An object is always conceived of as a particular located in space (and time) and characterized by a certain property. Thus, Kant's position comes close to Evans' position:

When we represent material objects in the imagination, we ipso facto represent them as located and differentiated in space. We imagine the carpenter, in the example of 4.2, as located at a particular position in space, though, of course, there is no particular position we imagine him as having. Such representations of



objects in the imagination are just like arbitrary fundamental Ideas (to be understood on the model of the arbitrary names of certain formal systems). (Evans 1982: 114)

Kant's position thus seems to be that the understanding cannot grasp particulars; it can only grasp things through concepts, i.e. through properties which can in principle characterize multiple things. It could not even hope to grasp an object through a complete concept, for even then there could be massive reduplication. Only sensibility and imagination allow us to be presented with particulars, not because of their receptivity, but because of their intuitive nature. Sensibility and Imagination are intuitive because they present things under the forms of space and time, and it is these forms that are responsible for the capacity to demonstratively identify particulars. They can do this because they are irreducibly oriented, directional, topological. On this reading, we can demonstratively identify a particular not because we are related to it in a certain way, but because we place it and ourselves in the same oriented and orientable space. The very act of pointing, of saying "here" and "there", does not depend on some causal or perceptual relation between objects, but rather on the fact that space is always already grasped as directional, and that we already grasp ourselves as irreducibly spatially situated.

This account makes sense of much of Kant's insistences, and ascribes to him a more refined theory of the distinct contribution of sensibility. But before passing on I would also like to point out that it harmonizes with the Critique of the Power of Judgment. We saw that the problem there was the relation between the particular and the general, and that this problem arises because we are divided between sensibility and understanding. This is so because the relevant feature of sensibility is that it allows us to refer to particulars, and the relevant feature of understanding is that it allows us to grasp through general properties. For this reason, whenever we think understanding yields us a concept of an object by itself, we are subject to illusion, because understanding cannot give us things; it can give us only properties.

Furthermore, this explains why the contribution of sensibility is authoritative. One of the problems presented by the Myth of the Given is that it is not clear why the contribution of sensibility should matter to the conceptual, i.e. why in our conceptual activity we should "care" about the empirical. If it is true that sensibility is primarily responsible for presenting us with particulars, then it is obvious from where this authority stems. As I argued in chapter 5, Kant believes that reason has an intrinsic interest in striving for the recognition of particularity. It would thus seem that Kant's sensibility matters to the understanding because it is sanctioned by reasons interest in specificity, which counteracts the interest in unity looked after by understanding. This will become clearer in the 7<sup>th</sup> and final chapter of this dissertation, in my discussion of the harmony of the faculties in aesthetic judgment. For now, I will simply conclude that Kant maintained the distinction of sensibility and understanding as a core doctrine of

his critical system, and that this does not subject him to the Myth of the Given, since the notionally distinct contribution of sensibility to thought is not its receptivity, but rather its intuitive nature, which is due to its a priori form, and not to its a posteriori matter.

### 6.3.2 The Paralogisms and the Unity of the Soul

In the previous subsection, I have tried to indicate why Kant believed there to be distinct and irreducible faculties, insofar as these faculties correspond with general moments of epistemic activity. I have argued that the lower faculties are indispensable and irreducible because they allow us to refer to particulars, a capacity that depends on our capacity to orient ourselves in a space-time of which we are in principle not the sole occupants. In admitting, however, that the soul, or at least its transcendental equivalent, consists of such a multiplicity of faculties, Kant seems to be opening himself up to a very simple charge, namely that of contradicting the simplicity of the soul. Wolff argued, after all, that the soul had to have a basic power and only one basic power because it is obviously a simple substance. It is *obviously* so, in the opinion of the German rationalists, because thought itself is simple and individuated in a way that only a true simple substance, and no aggregate can be. Crusius did not really challenge the assumption that the soul is a simple substance, but he did challenge the idea that the soul's harboring multiple irreducible fundamental powers would imply that it is composite. The problem, however, is that he seems to assume that the composition of mental forces is *sui generis* and at best analogous to the composition of physical forces, simply because they have to compose within a single simple substance.

In many ways, Kant was a follower of Crusius rather than of Wolff, and therefore defended some version of faculty pluralism. As a result, like Crusius, he is obliged to explain how the multiplicity of the faculties can be squared with the nature of the soul as it reveals itself in the very conception of thought. Such an argument can be found in the Paralogisms-chapter, where Kant targets the claims made by rational psychology and shows that they depend on illegitimate subreptions. In this subsection, I will argue that the Paralogisms, and concomitantly Kant's theory of the self, are not just intended to bar the substantial metaphysical claims made in the *Schulphilosophie*, but that they are also meant to offer an alternative understanding of the unity that characterizes thought. Specifically, I will argue that Kant wanted to challenge the claim that thought immediately reveals itself to have a substantial (analytic) unity. In the next section, I will argue that Kant instead wished to think of thought, or of epistemic activity, as exhibiting a normative, organic unity.

In discussing the Paralogisms chapter, it is important to note that it was all but completely rewritten for the second edition of the *Critique of Pure Reason*. Any interpretation must therefore offer some rationale for Kant's choice to rewrite it. In his

study of the philosophy of mind that supposedly lies behind Kant's philosophy, Andrew Brook (1994: 113-114) suggests that the substantial change in the Paralogisms chapter is at least correlated with one other chapter that was almost completely rewritten for the second edition: the transcendental deduction. He submits that Kant was forced to go into details about the nature of apperception in the Paralogisms chapter because he had not done so in the transcendental deduction, where he originally introduced the concept. In the B-deduction, Brook argues, we therefore find more of the comments that the first edition had postponed until the Transcendental Dialectic. I believe he is at least right in identifying the explanandum, namely the fact that the most valuable comments on the nature of transcendental apperception are to be found in the B-Deduction and the A-Paralogisms.

Kant's conception of transcendental apperception can only be understood in relation to his theory of synthesis. This is reflected by the structure of the completely rewritten second-edition version of the Transcendental Deduction proper, which opens with the following passage:

the combination (*conjunctio*) of a manifold in general can never come to us through the senses, and therefore cannot already be contained in the pure form of sensible intuition; for it is an act of the spontaneity of the power of representation, and, since one must call the latter understanding, in distinction from sensibility, all combination, whether we are conscious of it or not, whether it is a combination of the manifold of intuition or of several concepts, and in the first case either of sensible or non-sensible intuition, is an action of the understanding, which we would designate with the general title *synthesis* in order at the same time to draw attention to the fact that we can represent nothing as combined in the object without having previously combined it ourselves, and that among all representations *combination* is the only one that is not given through objects but can be executed only by the subject itself, since it is an act of its self-activity. One can here easily see that this action must originally be unitary and equally valid for all combination, and that the dissolution (*analysis*) that seems to be its opposite, in fact always presupposes it; for where the understanding has not previously combined anything, neither can it dissolve anything, for only through it can something have been given to the power of representation as combined.  
(CPR: B 129-130)

Kant makes two distinct though related claims here, namely the spontaneity of synthesis and the primacy of synthesis. The first claim is that sensibility, being receptive, can never give us a combination of a manifold, but only the manifold itself. Combinations are never simply given; they require a specific act on the part of the subject that Kant calls synthesis. This leads to the second claim, that of the primacy of synthesis. If we are never simply presented with a synthetic whole, than all synthetic wholes are due to synthetic activities. Kant makes the latter point by stating that all

analysis is preceded by a synthesis. He thus refuses any account of thought that regards thinking primarily as analyzing a given complex whole into its simple constituents. Such an account seems to lie behind many Early Modern epistemologies. Kant states that analysis would not be possible if there were not already some capacity for thinking synthetic wholes, a capacity that needs to be distinct from sensibility. These sparse comments refer, of course, to the theory of the subjective deduction, which opened the A-deduction, but is absent in the B-deduction.

After having introduced his theory that the very idea of a combination requires a capacity for synthesis, Kant continues §15 by noting that a further concept is required, namely that of unity:

But in addition to the concept of the manifold and of its synthesis, the concept of combination also carries with it the concept of the unity of the manifold. Combination is the representation of the *synthetic* unity of the manifold. The representation of this unity cannot, therefore, arise from the combination; rather, by being added to the representation of the manifold, it first makes the concept of combination possible. This unity, which precedes all concepts of combination a priori, is not the former category of unity (§ 10); for all categories are grounded on logical functions in judgments, but in these combinations, thus the unity of given concepts, is already thought. The category therefore already presupposes combination. We must therefore seek this unity (as qualitative, § 12) someplace higher, namely in that which itself contains the ground of the unity of different concepts in judgments, and hence of the possibility of the understanding, even in its logical use. (CPR: B 130-131)

This passage again announces two fundamental glosses on the kind of unity that is required for the capacity to be presented with combinations, and again they are importantly related. The first gloss is that the unity under consideration is synthetic, and, as Kant will indicate in the following paragraph, not analytic. We can already understand this by reminding ourselves of the already introduced doctrine of the primacy of synthesis: the analytic unity is the unity that can be analyzed into its constituents. But this unity already presupposes a synthetic activity that has accomplished the unity, as Kant states in a footnote two pages later: “only by means of an antecedently conceived possible synthetic unity can I represent to myself the analytical unity” (CPR: B 133). The second gloss on the synthetic unity is that it is not to be conflated with the kind of unity that is thought under the category of unity. Kant claims that, in fact, the very capacity to think something under the category of unity depends on the accomplishment of the more fundamental synthetic unity.

Having introduced the idea of an original synthetic unity, Kant goes on to characterize this unity in §17, which is aptly titled “On the original-synthetic unity of apperception”, by connecting it with the idea of being represented in the same subject:

all manifold of intuition has a necessary relation to the *I think* in the same subject in which this manifold is to be encountered. But this representation is an act of *spontaneity*, i.e., it cannot be regarded as belonging to sensibility. I call it the *pure apperception*, in order to distinguish it from the *empirical* one, or also the *original apperception*, since it is that self-consciousness which, because it produces the representation *I think*, which must be able to accompany all others and which in all consciousness is one and the same, cannot be accompanied by any further representation. I also call its unity the *transcendental* unity of self-consciousness in order to designate the possibility of a priori cognition from it. For the manifold representations that are given in a certain intuition would not all together be *my* representations if they did not all together belong to a self-consciousness; i.e., as *my* representations (even if I am not conscious of them as such) they must yet necessarily be in accord with the condition under which alone they can stand together in a universal self-consciousness, because otherwise they would not throughout belong to me. (CPR: B 131-133)

Kant thus claims that the original unity of synthesis is a unity in the sense that the manifold unified in it is ascribed to the same subject, and therefore connected in the same thought. It is easy to misunderstand this as the theory that, in representing and knowing, we are capable of relating thoughts because they are all somehow modes or attributes inhering in the same substance or subject. However, this latter view, which is a version of Cartesianism popular among early modern rationalists, is precisely the target of Kant's criticism. Kant's view is not that thoughts are connected by being in the same subject, it is that they are understood as being in the same subject. But in order to be in the same subject in the latter way, in order to exhibit the latter unity, they need to first be unified through an original synthesis:

this thoroughgoing identity of the apperception of a manifold given in intuition contains a synthesis of the representations, and is possible only through the consciousness of this synthesis. For the empirical consciousness that accompanies different representations is by itself dispersed and without relation to the identity of the subject. The latter relation therefore does not yet come about by my accompanying each representation with consciousness, but rather by my adding one representation to the other and being conscious of their synthesis. Therefore it is only because I can combine a manifold of given representations in one consciousness that it is possible for me to represent the identity of the consciousness in these representations itself, i.e., the analytical unity of apperception is only possible under the presupposition of some synthetic one. (CPR: B 133)

I believe Kant can best be understood here as claiming that the manifold of representations is not itself a unity in the sense of belonging to a single subject. To exhibit the unity of the subject, the manifold must be unified by a specific original

operation, which is active, but which also precedes any conscious or intentional activity, since conscious activity is only possible on the condition that this original unification has already been effected. But then which function is responsible for this unity? According to Kant, this function is the transcendental apperception, the original unity, and interestingly this transcendental apperception forms the core of the faculty of understanding. This is apparent from Kant's statement, in the A-deduction that "[t]he unity of apperception in relation to the synthesis of the imagination is the understanding, and this very same unity, in relation to the transcendental synthesis of the imagination, is the pure understanding" (CPR A 119). This statement suggests that the understanding is the working of transcendental apperception on the materials of sensibility and understanding. This makes sense if we understand Kant's definition of understanding as a "faculty of rules" (CPR A 126):

The understanding is thus not merely a faculty for making rules through the comparison of the appearances; it is itself the legislation for nature, i.e., without understanding there would not be any nature at all, i.e., synthetic unity of the manifold of appearances in accordance with rules; for appearances, as such, cannot occur outside us, exist only in our sensibility. The latter, however, as the object of cognition in an experience, with everything it may contain, is possible only in the unity of apperception. The unity of apperception, however, is the transcendental ground of the necessary lawfulness of all appearances in an experience. This very same unity of apperception with regard to a manifold of representations (that namely of determining it out of a single one) is the rule, and the faculty of these rules is the understanding. (CPR A 126-127)

The understanding, as a faculty of unifying and synthesizing under rules, has as a core the transcendental ground of this unity: the transcendental apperception. If this is true, and if it is also the case, as Kant claims, that sensibility itself does not exhibit this unity, then the unity of the subject is not a pre-given – it is an accomplishment by one of the faculties as it works in concert with others, in this case sensibility and imagination.

Such an interpretation provides a valuable clue for identifying how the arguments of the paralogisms, which are criticisms of the metaphysical conclusions drawn from the nature of selfhood, pertain to the debate on the doctrine of the faculties. Indeed, I will now try to show that one line of argument in the first two A-paralogisms is that rational psychologists like Wolff conflate the synthetic and the analytic unity of the subject, and therefore interpret what is, for Kant, an underlying unifying function with an underlying unity.

Although I will not be offering a full interpretation of the paralogisms here, I will start off by discussing the notion of a paralogism. Kant opens the chapter as follows:

A logical paralogism consists in the falsity of a syllogism due to its form, whatever its content may otherwise be. A transcendental paralogism, however, has a

transcendental ground for inferring falsely due to its form. Thus a fallacy of this kind will have its ground in the nature of human reason, and will bring with it an unavoidable, although not insoluble, illusion. (CPR A 341 / B 399)

He later tries to clarify in what sense the paralogisms are logical fallacies by specifying the kind of fallacy involved:

If one wants to give a logical title to the paralogism in the dialectical syllogisms of the rational doctrine of the soul, insofar as they have correct premises, then it can count as *sophisma figurae dictionis*, in which the major premise makes a merely transcendental use of the category, in regard to its condition, but in which the minor premise and the conclusion, in respect of the soul that is subsumed under this condition, make an empirical use of the same category. (CPR A 402-403)

Indeed, Kant seemed to have regarded the *sophisma figurae dictionis* as a paradigmatic example of a paralogism, for instance in the *Jäsche Logik*:

An inference of reason that is wrong as to form, although it has for itself the illusion of a correct inference, is called a fallacy (fallacia). Such an inference is a paralogism insofar as one deceives oneself through it, a sophism insofar as one intentionally seeks to deceive others through it.

Note. The ancients occupied themselves very much with the art of making such sophisms. Therefore many of this kind have emerged, e.g., the *sophisma figurae dictionis*, in which the *medius terminus* is taken in different meanings - fallacia a dicta secundum quid ad dictum simpliciter, *sophisma heterozeteseos*, *elenchid ignorationis*, etc. (AA IX: 134-135)

Here, the fallacy of *figurae dictionis* is explained as a fallacy through the equivocation of the middle term. A syllogism is meant to connect the extreme terms through a third term, a middle term, which occurs in both the minor and the major proposition. But as Kant notes, Medieval logicians have often warned that the syllogism might be invalid if there is equivocation of the Middle term, i.e. if the middle term is taken in different ways in the different premisses. To us, this might not sound like a formal invalidity, but rather like a material invalidity. But we should not overlook one crucial feature of the *figurae dictionis* that would not have been overlooked by Kant, namely that this fallacy often depends on *category mistakes*. This is clear from the longer treatment of the fallacy in the *Hecschel Logic*:

The *sophisma figurae dictionis* is where the *medius terminus* is taken in different meanings. E.g., A philosopher is a kind of learned man. Leibniz was a philosopher, consequently Leibniz was a genus of learned man. *Vox medii termini, philosopher*, is taken in different meanings[;] one time it is taken as a predicate, and the other time as a multitude of things to which the predicate belongs. When the *medius terminus* in the two premises is taken in different meanings, then this always

yields a *fallacium*. E.g., no artist is born; some men are artists[;] hence some men are not born. In the *major* the *medius terminus* means the art, and in the *minor* the man. (Kant 1992: 458)

To understand how equivocation constitutes a logical fallacy, and not just a confusion, we must return to the context of homonymy and synonymy in Aristotelian logic. In the very first sentence of the *Categories*, Aristotle defines homonymy: “When things have only a name in common and the definition of being which corresponds to the name is different, they are called homonymous. Thus, for example, both a man and a picture are animals.” (1a1) The equivocation here is between the man and the depicted man: the former is a man truly, the latter is a representation of a man. The fallacy is based on the following rule: if two things are of the same species, they are of the same genus, and of the same higher genera as well. Yet, although the word man is equally applied to a real man and to a depicted man, these two things are not truly of the same genus. In the case of equivocation, two things are called the same because they share some difference, some distinct mark, although these differences or marks are not predicated of the same genera. In Aristotle’s case, both things are called man because they are man-shaped, but the former is a man-shaped animal and the latter is a man-shaped picture.

This discussion explains two peculiarities about Kant’s conception of the paralogisms. First of all, it explains why Kant thought that the fallacies were formal errors: as we saw in 5.3.1, the semantics of species and genera forms an integral part of his logic, and is therefore properly understood as logical. Because the terms in logical propositions stand for spaces within logical space, it is a logical error to conflate such spaces, rather than an empirical or linguistic one. When Kant introduces the concept of a *transcendental* paralogism, he does not mean to introduce a fallacy that is not purely logical, but rather a form of purely logical fallacy where the equivocation is unavoidable because it is suggested by the nature of human reason.

The second peculiarity we can explain is precisely why the employment of a category would be a necessary equivocation. As I have noted, a case of homonymy is usually due to some similarity between two things that are not of the same genus. There is in the Aristotelian tradition a strand of reflection on what it would mean if the two did not even fall under the same “highest genus”, i.e. under the same category. In such a case, the term would be applied homonymously to the two things because they are not of the same category. According to Aristotle, being is precisely such a term, i.e. a term that can only be applied across categories through equivocation (in fact, the Aristotelian categories are best understood as the many different senses of the fundamentally equivocal term “to be”, as is explained in *Metaphysics* IV.2). For this reason, being is often called a transcendental, since it transcends the categories.

It is interesting that Kant himself mentions the notion of a transcendental term in the so-called “*Metaphysical Deduction*”. There, he reminds us of the Medieval theory



that “unum, verum et bonum” are three convertible senses of the transcendental term “being”, and objects to it:

These supposedly transcendental predicates of things are nothing other than logical requisites and criteria of all cognition of things in general, and ground it in the categories of quantity, namely, the categories of unity, plurality, and totality; yet these categories must really have been taken as material, as belonging to the possibility of things itself, when in fact they should have been used in a merely formal sense, as belonging to the logical requirements for every cognition; thus these criteria of thinking were carelessly made into properties of things in themselves. In every cognition of an object there is, namely, unity of the concept, which one can call qualitative unity insofar as by that only the unity of the comprehension of the manifold of cognition is thought, as, say, the unity of the theme in a play, a speech, or a fable. Second, truth in respect of the consequences. The more true consequences from a given concept, the more indication of its objective reality. One could call this the qualitative plurality of the marks that belong to a concept as a common ground (not of in it as a magnitude). Third, finally, perfection, which consists in this plurality conversely being traced back to the unity of the concept, and agreeing completely with this one and no other one, which one can call qualitative completeness (totality). From this it is obvious these logical criteria of the possibility of cognition in general transform the three categories of magnitude, in which the unity in the generation of the magnitude must be assumed to be completely homogeneous, into a principle with the quality of a cognition for the connection of heterogeneous elements of cognition into one consciousness also. (CPR B 114-115)

In this passage, which was added in the second edition, Kant tells us that the material use made of these transcendental concepts is in fact a usage which is within the category of quantity rather than transcendent to it. After all, Kant believes that metaphysics cannot meaningfully speak of that which transcends the categories. That is why he must object to the material, rather than merely logical, use of the categories of quantity. In this case, he seems most concerned with the case of unity, which is here fallaciously used in order to provide a concept of the unity of heterogeneous, rather than that of merely homogeneous natures. This comment pertains directly to the paralogisms, as we will see.

Now that we have a better conception of a paralogism in general and of the transcendental paralogisms in particular, we can look at the first two paralogisms and see what makes them transcendently paralogistic. The first paralogism runs as follows:

That the representation of which is the absolute subject of our judgments, and hence cannot be used as the determination of another thing, is substance.

I, as a thinking being, am the absolute subject of all my possible judgments and this representation Myself cannot be used as the predicate of any other thing. Thus I, as thinking being (soul), am substance. (CPR A 348)

Kant opens his discussion of the paralogism with the following passage, which does not itself constitute a criticism, but rather sets up his treatment of the problem:

We have shown in the analytical part of the Transcendental Logic that pure categories (and among them also the category of substance) have in themselves no objective significance at all unless an intuition is subsumed under them, to the manifold of which they can be applied as functions of synthetic unity. Without that they are merely functions of a judgment without content. Of any thing in general I can say that it is a substance, insofar as I distinguish it from mere predicates and determinations of things. Now in all our thinking the I is the subject, in which thoughts inhere only as determinations, and this I cannot be used as the determination of another thing. Thus everyone must necessarily regard Himself as a substance, but regard his thinking only as accidents of his existence and determinations of his state. (CPR A 348-349)

The first half of this passage reminds us of the constraint on the usage of the category of substance, namely that it can only legitimately be used when applied to an intuition. Kant states that without such a sensible element, the categories are merely functions of judgment without content. However, Kant does say that we are allowed to use the category in such cases, provided we use it in a purely logical way. I believe we can understand this better if we keep in mind that there have traditionally been two ways of using the concept of a substance, namely firstly as a proper subject of predication, and secondly as an independently existing thing. For Aristotle, both concepts were so intimately entwined that he believed that to be a metaphysically independent thing meant to be a proper subject of predication and conversely. If we read Kant as seeking to distinguish those two meanings of the term, we can read him as saying that we are systematically tempted to regard ourselves as the metaphysical substance underlying our thoughts because we cannot but regard ourselves as the logical subject of all our thoughts. After all, the transcendental unity requires that we regard all thoughts as connected in the same subject, i.e. as predicated of the same subject. But this does not of itself imply that all thoughts are the attributes of the same metaphysical substance. That the latter inference is problematic, appears from the continuation of the discussion of the first paralogism:

But now what sort of use am I to make of this concept of a substance? That I, as a thinking being, endured for myself, that naturally I neither arise nor perish - this I can by no means infer, and yet it is for that alone that the concept of the substantiality of my thinking subject can be useful to me; without that I could very well dispense with it altogether.

So much is lacking for us to be able to infer these properties solely from the pure category of substance, that we must rather ground the persistence of a given object on experience if we would apply to that object the empirically usable concept of a substance. But now we have not grounded the present proposition on any experience, but have merely inferred [it] from the concept of the relation that all thought has to the I as the common subject in which it inheres. Nor would we be able to establish such a persistence through any secure observation, even if we supposed one. For the I is, to be sure, in all thoughts; but not the least intuition is bound up with this representation, which would distinguish it from other objects of intuition. Therefore one can, to be sure, perceive that this representation continually recurs with every thought, but not that it is a standing and abiding intuition, in which thoughts (as variable) would change. (CPR A 350-351)

The first paragraph lists some of the implications that metaphysicians have sought to draw from the fact that we must appear to ourselves as the substances in which our thoughts inhere. Kant immediately repudiates these implications, stating that they do not follow from the category of substance in itself, but rather from the category of substance as applied to empirical objects. Again, this statement cannot be understood unless we take the term “substance” to be equivocal here, namely in the sense of either the pure or of the schematized category. The pure category is limited to the idea of a subject of predication, whereas the schematized category means “persistent unity underlying change”. Since our self is not really intuited by us, but rather the result of a necessary function of the *Gemüt*, or of the understanding, we cannot take it as an empirical basis. Hence, we are not allowed to apply the category of substance to it in more than the purely formal sense.

Kant’s whole discussion here is muddled, however, which may be an important reason why he chose to rewrite it for the second edition. The major problem is that, thus far, Kant locates the equivocation in the term substance. This cannot be the ground for the paralogistic nature of the argument, however, since the term substance is not even the middle term; the term “absolute subject of our judgments” is. I have already indicated how this confusion can be eliminated, namely by taking Kant to himself seemingly equivocate substance and subject, an equivocation that is understandable because of contemporary usage.

That Kant himself takes the paralogism to be an equivocation on the concept of subject, is immediately clear from the conclusion of his discussion in the first edition:

From this it follows that the first syllogism of transcendental psychology imposes on us an only allegedly new insight when it passes off the constant logical subject of thinking as the cognition of a real subject of inherence, with which we do not and cannot have the least acquaintance, because consciousness is the one single thing that makes all representations into thoughts, and in which, therefore, as in

the transcendental subject, our perceptions must be encountered; and apart from this logical significance of the I, we have no acquaintance with the subject in itself that grounds this I as a substratum, just as it grounds thoughts. (CPR A 350)

Here, Kant explicitly states that the equivocation is between the purely logical concept of a subject as a subject of predication and the metaphysical concept of a subject as a subject of inherence. The same seems to lie behind the B-edition version of the first paralogism, which Kant rebuts as follows:

The major premise talks about a being that can be thought of in every respect, and consequently even as it might be given in intuition. But the minor premise talks about this being only insofar as it is considered as subject, relative only to thinking and the unity of consciousness, but not at the same time in relation to the intuition through which it is given as an object for thinking. Thus the conclusion is drawn *per Sophisma figurae dictionis*, hence means of a deceptive inference. (CPR B 411)

Here the equivocation appears to be between the different ways in which we take something to be the subject of something. Kant clarifies in a footnote that the equivocation is due to the fact that the relation between thought and its subject is simply not of the same kind as the relation between a substance and its properties:

"Thinking" is taken in an entirely different signification in the two premises: in the major premise, as it applies to an object in general (hence as it may be given in intuition); but in the minor premise only as it subsists in relation to self-consciousness, where, therefore, no object is thought, but only the relation to oneself as subject (as the form of thinking) is represented. In the first premise, things are talked about that cannot be thought of other than as subjects; the second premise, however, talks not about things, but about thinking (in that one abstracts from every object), in which the I always serves as subject of consciousness; hence in the conclusion it cannot follow that I cannot exist otherwise than as subject, but rather only that in thinking my existence I can use myself only as the subject of judgment, which is an identical proposition, that discloses absolutely nothing about the manner of my existence. (CPR B 411-412)

Kant again uses substance and subject in a confusing way here, since he says that "it cannot follow that I cannot exist otherwise than as subject, but rather only that in thinking my existence I can use myself only as the subject of judgment". We would expect him to say that we cannot think of ourselves as substance, although we can think of ourselves as substance. On closer inspection, however, the passage is less problematic: he is talking about that which cannot exist otherwise than as subject, and that is precisely the definition of substance. The proper reading then becomes: although I am allowed to think of myself as the subject of thinking, I am not allowed to think of myself as the thinking existent. We can easily see why this is so on the basis of our

discussion of the transcendental apperception: the subject of thought is not itself an object, it is rather the mark of a function of thought, the unifying function. That this underlies the argument of the paralogisms is even clearer in the case of the second paralogism.

The second paralogism is a purported argument for the simplicity of the soul. Kant presents the argument as follows in the first edition of the *Critique of Pure Reason*:

That thing whose action can never be regarded as the concurrence of many acting things, is simple.

Now the soul, or the thinking I, is such a thing.

Thus etc. (CPR A 351)

As was the case with the first paralogism, here too the syllogism Kant offers does not easily reveal the structure and fallacious nature of the argument. Consequently, we must again look at his discussion of the argument to gain insight into these aspects:

The so-called *nervus probandi* of this argument lies in the proposition that many representations have to be contained in the absolute unity of the thinking subject in order to constitute one thought. But no one can prove this proposition from concepts. For how could he set about to accomplish this? The proposition "A thought can be only the effect of the absolute unity of a thinking being" cannot be treated as analytic. For the unity of a thought consisting of many representations is collective, and, as far as mere concepts are concerned, it can be related to the collective unity of the substances cooperating in it (as the movement of a body is the composite movement of all its parts) just as easily as to the absolute unity of the subject. Thus there can be no insight into the necessity of presupposing a simple substance for a composite thought according to the rule of identity. But that this same proposition should be cognized synthetically and fully *a priori* from sheer concepts - that answer no one will trust himself to give when he has insight into the ground of the possibility of synthetic propositions as we have established it above. (CPR A 352-353)

Kant interestingly remarks here that the idea of the unity of thought does not analytically entail the absolute unity, i.e. simplicity, of the thing in which thought inheres. He argues for this by remarking that thought, though it is a unity, can (and does) consist of a multiplicity of representations that are united in one whole. He then suggests that this collective unity can be just as easily due to the cooperation of the different substances bearing those different representations as it can be due to the fact that they all inhere in the same simple substance. This first stage of Kant's argument is interesting in light of the debate on the unity of the faculties. Both Wolff and Crusius adhered to the theory that the soul is immaterial because it exhibits a specific kind of unity. Yet, Wolff deduced from this that, being a simple substance, the soul must also be simple in terms of its fundamental forces. Crusius argued instead that, although the soul

is immaterial and simple, it can nevertheless be the bearer of multiple forces. These forces compose in the way that the movements of the different parts of an extended body compose. I therefore submit that Kant is attacking the doctrine of the simplicity of the soul not in order to disprove immaterialism (which he arguably himself prefers as a position), but in order to make room for the possibility that thought, although it exhibits a characteristic unity, can nevertheless be due to the cooperation of a multiplicity of forces, rather than due to a single unitary principle.

Further evidence for this is the long discussion on the connection between immateriality and simplicity Kant appends to his treatment of the second paralogism. There, he first notes that the simplicity of the soul is usually adduced as a premise for the argument that the soul is immaterial (CPR A 356). He then notes that, even if we were to concede that premise, the conclusion would not follow (CPR A 357). I will abstain from offering a full interpretation of the argument for this, because it would lead us too far into the murky swamp that is the debate on the nature and value of transcendental idealism. It is in itself important, however, that Kant seeks to dissociate his argument for the simplicity of the soul from that of its immateriality. This may mean that his purported adherence to immaterialism does not condemn him to accept simplicity as well.

Kant dismisses the argument for the simplicity of the soul on the familiar ground that it includes an equivocation on the unity of thought:

Here, therefore, as in the previous paralogism, the formal proposition of apperception, I think, remains the entire ground on which rational psychology ventures to extend its cognitions; this proposition is of course obviously not an experience, but rather the form of apperception, on which every experience depends and which precedes it, yet it must nevertheless always be regarded only in regard to a possible cognition in general, as its merely subjective condition, which we unjustly make into a condition of possibility of the cognition of objects, namely into a concept of a thinking being in general, because we are unable to represent this being without positing ourselves along with the formula of our consciousness, in place of every other intelligent being. (CPR A 354)

Here, the equivocation is yet again between the transcendental apperception as a function of thought and the soul as the thinking thing. The former exhibits a formal unity, which is a condition for the possibility of thought, whereas the latter is supposed to be a substantial unity, a simplicity, and thus a property of an object of thought. Kant's fundamental point of the paralogisms is thus that, although the apperception is necessary for the possibility of thought, it cannot itself be the object of thought. We can understand this better by returning to the debate on the concept of a transcendental. As I explained above, Kant criticized metaphysicians for making material and constitutive use of terms that supposedly transcend the categories. His transcendental deduction is meant as an argument that material use can only be made of concepts *under* the

categories, and then only when these categories are applied to intuitions, i.e. to sensibly mediated objects. So, instead of dealing with what transcends the categories, metaphysicians actually apply the categories in cases where they are not applicable. This does not mean that Kant refuses the idea of that which “transcends” the categories altogether; on the contrary, the transcendental apperception is indeed such a category-transcending concept. But it is not transcendent to the categories because it metaphysically underlies them – it transcends them because it first makes them possible. From this perspective, we can understand Kant’s insistence in the B-deduction that the unity of the apperception should not be conflated with the kind of unity thought under the category of unity, since the latter is only applicable if the former has already been effected. The unity of thought is a function of thought, not the structure of the thinking thing (*res cogitans*), and our thinking is a unity to us because a distinct function has unified our representations, not because it was always already a single simple substance. And yet, Kant believes that we are constantly seduced to this conflation, which needn’t be harmful of itself. In the next subsection we will see which positive use can be made of the concept of the soul.

### 6.3.3 The Disunity of the Faculties and the Regulative Ideas

It is well known that Kant, for all his misgivings regarding the nature of the soul and the possibility of meaningfully speaking of it, thought that the concept has definite use in philosophy. Usually, however, we are directed to his practical philosophy, where the postulate of the immortality of the soul plays a role, for this positive employment of the idea. This is not entirely adequate, since Kant attributes to the idea of the soul a legitimate theoretical use as well, specifically in the Appendix to the Transcendental Dialectic.

As we saw in chapter 5, Kant believes that reason, although it leads us into illusion and temptation, nevertheless has a legitimate employment in theoretical philosophy, through the idea of systematicity. It is for the same reason that he accords the idea of the soul a legitimate employment:

The first object of such an idea is myself considered as thinking nature (soul). If I want to seek out the properties with which a thinking thing exists in itself, then I have to ask experience, and I cannot even apply any of the categories to this object except insofar as its schema is given in sensible intuition. By this means, however, I will never attain to a systematic unity of all the appearances of inner sense. Thus instead of the concept of experience (of that which the soul actually is), which cannot lead us very far, reason takes the concept of the empirical unity of all thought, and, by thinking this unity unconditionally and originally, it makes out of it a concept of reason (an idea) of a simple substance, unchangeable in itself

(identical in personality), standing in community with other real things outside it – in a word, the concept of a simple self-sufficient intelligence. With this, however, reason has nothing before its eyes except a principle of the systematic unity in explaining the appearances of the soul, namely by considering all determinations as in one subject, all powers, as far as possible, as derived from one unique fundamental power, all change as belonging to the states of one and the same persisting being, and by representing all appearances in space as entirely distinct from the acts of thinking. That simplicity of substance, etc., ought to be only the schema for this regulative principle, and it is not presupposed as if it were the real ground of properties of the soul. For these properties could rest on entirely different grounds, with which we are not acquainted at all, just as we might not really be able to cognize the soul at all through these assumed predicates even if we let them hold of it absolutely, since they constitute a mere idea that cannot be represented in concreto at all. Now nothing but advantage can arise from such a psychological idea, if only one guards against letting it hold as something more than a mere idea, i.e., if one lets it hold merely relative to the systematic use of reason in respect of the appearances of our soul. (CPR A 682-683 / B 710-711)

I will focus specifically on Kant's claim that the idea of the simplicity of the soul, when taken merely regulatively, can aid us in the investigation of the powers of the soul. It would do so by supposing that the many powers of the soul are all in some way specifications of a single fundamental power, which would add systematicity to the idea of the soul. This theory can help us understand Kant's relation to Wolff, Crusius and Tetens respectively. Kant criticizes Wolff for using the idea of the simplicity of the soul determinately, thereby merely postulating that the many powers of the soul must ultimately reduce to a single fundamental power. No such determinate claims can be made, and hence it cannot be ruled out that the many powers of the soul relate in a different manner. Yet, by allowing the regulative use of the idea, Kant is equally criticizing Crusius' approach, which allowed for the multiplicity of the powers. This risks terminating in a proliferation of powers, all of which are qualitatively distinct, and perhaps even the postulation of different powers for each different representation in the mind. Kant realizes that this is unsatisfactory, and probably realizes that Crusius would agree with this. He therefore encourages Tetens' attempt to reduce the powers of the mind as far as possible without doing injustice to the phenomena.<sup>10</sup>

Some care is to be taken in interpreting this principle in Kant, for it has led many to believe that, deep down, Kant adhered to the doctrine of the unique fundamental force, even though he did not believe it could be proven. The Hegelian and Heideggerian

---

<sup>10</sup> Hessbrüggen-Walter (2004: 161) has also suggested that this passage is directly addressed to Tetens.



readings discussed in the first section of this chapter are precisely such attempts of ascribing to Kant a determinate employment of the idea of a fundamental force. This would miss, however, the importance of the shift of context: Kant is no longer discussing unity as such, but systematic unity. Earlier in the Appendix, Kant had already introduced the debate on the unity and disunity of the faculties as an example of a debate where systematicity operates:

Among the different kinds of unity according to concepts of the understanding belongs the causality of a substance, which is called "power." At first glance the various appearances of one and the same substance show such diversity that one must assume almost as many powers as there are effects, as in the human mind there are sensation, consciousness, imagination, memory, wit, the power to distinguish, pleasure, desire, etc. Initially a logical maxim bids us to reduce this apparent variety as far as possible by discovering hidden identity through comparison, and seeing if imagination combined with consciousness may not be memory, wit, the power to distinguish, or perhaps even understanding and reason. The idea of a fundamental power - though logic does not at all ascertain whether there is such a thing - is at least the problem set by a systematic representation of the manifoldness of powers. The logical principle of reason demands this unity as far as it is possible to bring it about, and the more appearances of this power and that power are found to be identical, the more probable it becomes that they are nothing but various expressions of one and the same power, which can be called (comparatively) their fundamental power. One proceeds in just the same way with the rest of the powers.

These comparatively fundamental powers must once again be compared with one another, so as to discover their unanimity and thereby bring them close to a single radical, i.e., absolutely fundamental, power. But this unity of reason is merely hypothetical. One asserts not that such a power must in fact be found, but rather that one must seek it for the benefit of reason, namely for setting up certain principles for the many rules with which experience may furnish us, and that where it can be done, one must in such a way bring systematic unity into cognition. (CPR A 649-650 / B 677-678)

Kant acknowledges that the striving for unity is indeed demanded by the systematic drive of reason, and that we are therefore rightly and understandably engaged in systematizing the various powers of the mind. It thus seems that the idea of an absolutely fundamental power is a regulative idea of reason in its theoretical employment. But when we look to the passage to see for which science it is a regulative idea, we find that it is *logic*, since Kant states that it is up to logic to ascertain whether there is such a thing as a fundamental power. This comment can admittedly be read in two ways. On the first reading, Kant is saying that the idea of an absolute power is dictated by the logical use of reason, but the logical use itself cannot determinately (and only regulatively) posit such an absolute power. Although this reading makes some

sense, it squares less with the text than another. On this reading, the logical use of reason sets, within each science, a regulative ideal, which this science itself cannot reach, i.e. determinately use or prove. In logic, the idea of absolutely fundamental power is such an idea. Although this idea gives guidance to the proper systematization and classification of epistemic functions and activities, it cannot be proven as such and can never be determinately used.

If this latter reading is correct, we see the following peculiarity: for Kant, the doctrine of the faculties is not the business of psychology, but rather that of *logic*, i.e. not of a descriptive science but of a normative science. Hessbrüggen-Walter (2004: 185) has shown that, in Kant's many writings, the doctrine of faculties is often a topic of logic rather than of psychology, and that in this respect Kant is in line with the tradition. The reason for this is that, for example, the three higher or discursive faculties correspond with the three major divisions of Early Modern textbooks on logic, which were thought to correspond with works from Aristotle's *Organon*.

Traditionally, *Categories*, *On Interpretation* and the *Analytics* were taken to deal with single terms, propositions and syllogisms respectively. Early Modern textbooks like the Port-Royal logic therefore usually consist of a part on terms, one on propositions and one on syllogisms. As Hessbrüggen-Walter (2004: 194-200) indicates, the same division is to be found in German textbooks on logic, and there the division corresponds with a division between three mental operations: the grasping of a single term by the understanding, the relating of two terms by the faculty of judgment and the relating of judgments and propositions through a syllogism by reason. Kant's tripartite division of the higher faculties neatly corresponds with these three mental operations:

General logic is constructed on a plan that corresponds quite precisely with the division of the higher faculties of cognition. These are: understanding, the power of judgment, and reason. In its analytic that doctrine accordingly deals with concepts, judgments, and inferences, corresponding exactly to the functions and the order of those powers of mind, which are comprehended under the broad designation of understanding in general. (*CPR* A 130-131 / B 169)

If the faculties are indeed best understood not as psychological properties, but rather as epistemic functions, then we can understand why Kant deals with them more as a topic of epistemology than as one of psychology. Moreover, it bears witness to a certain anti-psychologism in Kant. To see why this is so, just observe that anti-psychologistic authors of the late 20<sup>th</sup> century like Cohen, Frege and Husserl all used the word "logic" to refer to what we would now call pure epistemology, even though only Frege actively engaged in what we would now consider logic proper. This is because they found the term "epistemology" or "theory of knowledge" irreparably tainted by the psychologism of their predecessors. Logic being considered a normative rather than a descriptive science, was found to be a better term, even though it was obviously used in a broader

sense than the one usually accorded to it. The same move is recognizable in Kant, who had a word for the extended logic with which he was engaged, namely transcendental logic (in contrast with formal logic).<sup>11</sup>

This shows that the doctrine of faculties is improperly taken as a psychological topic in Kant; it is actually a logical and epistemological topic, concerned with the basic kinds of epistemic functions. Logic and epistemology are guided by a regulative ideal in their search for such basic kinds, namely that of an absolutely fundamental power. Yet, Kant warns us that this idea cannot be used constitutively, as determining an object, and hence warns against employing the metaphysical idea of a soul as a simple substance in order to legitimate this constitutive usage:

[I]f one attends to the transcendental use of the understanding, it is evident that this idea of a fundamental power in general does not function merely as a problem for hypothetical use, but pretends to objective reality, so that the systematic unity of a substance's many powers are postulated and an apodictic principle of reason is erected. For even without our having attempted to find the unanimity among the many powers, or indeed even when all such attempts to discover it have failed, we nevertheless presuppose that such a thing will be found; and it is not only, as in the case cited, on account of the unity of substance that reason presupposes systematic unity among the manifold powers, but rather reason does so even where many powers, though to a certain degree of the same kind, are found, as with matter in general, where particular natural laws stand under more general ones; and the parsimony of principles is not merely a principle of economy for reason, but becomes an inner law of its nature. (*CPR* A 650 / B 678)

Kant is obviously referring to the Wolffian position when he speaks of the presupposition of the unity of the faculties as required by the unity of the substance that is the soul. As we saw in the previous subsection, he denied that we could determinately claim of the soul that it is such a unity. But here, he says that the idea nevertheless has a legitimate employment, namely as a regulative ideal for the investigation of the faculties. Does this mean that Kant, in spite of his many misgivings regarding the premisses of such a position, ultimately does prefer and embrace he theory of the unity of the faculties? This would be a precipitous conclusion, for two reasons. First of all, Kant does not claim that unification is the only goal of systematicity, and secondly, he claims that the project of unification is a failure.

---

<sup>11</sup> This does not mean that there can be no reasonable questions regarding what such an anti-psychologistic theory of the faculties would look like, and whether it is a coherent position. These questions, however, lie beyond the scope of this dissertation.

As regards the first reason, we need to remind the reader of the conclusions of the previous chapter. There, we saw that Kant's concept of systematicity is not guided solely by the demand for unity, to the detriment of variety. In the context of the soul, this is easily forgotten, for the simple reason that the idea of the soul is, according to Kant, suggested by the idea of absolute unity (*CPR* A 334 / B 391). I interpret this as meaning that, as far as the soul is concerned, there is an internal metaphysical tendency of reason to regard it predominantly under the idea of absolute unity. But this would involve a culpable negligence towards the idea of variety. For Kant, we need to focus as much on the variety as on the unity of the faculties, if we want to do justice to the demand of reason. As Heßbrüggen-Walter (2004: 225) has noted, the two maxims are clearly balanced in Kant's reflections on the faculties, where unity is strived for, but the distinctness of several faculties is nevertheless maintained.

One important place where Kant rejects the idea of the unity of the faculties is in the first Introduction to the *Critique of the Power of Judgment*:

We can trace all faculties of the human mind without exception back to these three: the **faculty of cognition**, the **feeling of pleasure and displeasure**, and the **faculty of desire**. To be sure, philosophers who otherwise deserve nothing but praise for the thoroughness of their way of thinking have sought to explain this distinction as merely illusory and to reduce all faculties to the mere faculty of cognition. But it can easily be demonstrated, and has already been understood for some time, that this attempt to bring unity into the multiplicity of faculties, although undertaken in a genuinely philosophical spirit, is futile. (AA XX: 205-206)

Although this passage contains an explicit rejection of the idea of the unity of the faculties, which squares with the fact that Kant never explicitly adopts that idea except regulatively, it does raise further issues. After all, Kant is not speaking of any of the familiar faculties, namely sensibility, imagination, understanding, judgement and reason. Are we to understand this passage as claiming that the faculties are not unified, but that the latter five faculties are all reducible to the three faculties mentioned in the quote? Here again, a terminological peculiarity of Kant's system threatens to mislead us, since the word "faculty" is used equivocally here. Kant discerned two meanings of the word faculty. The first is the familiar one from our discussions, namely that of an epistemic function. The two lower and the three higher faculties are all such faculties. The second sense of faculty is the one invoked in the quoted passage, and is best understood as a specific orientation and internal alignment of all the faculties taken together. In a table offered by Kant at the end of the published introduction to the *Critique of the Power of Judgment*, these two "kinds" of faculties are represented as orthogonal to each other. This provides a clue for another table, namely that of all the faculties:

<i>faculty of cognition</i>	<i>feeling of pleasure and displeasure</i>	<i>faculty of desire</i>
Sensibility	Sensibility	Sensibility
Imagination	Imagination	Imagination
<b>Understanding</b>	Understanding	Understanding
Judgment	<b>Judgment</b>	Judgment
Reason	Reason	<b>Reason</b>

In this table, the faculties in the second sense are represented above the double line. Each faculty is an employment of the whole soul or *Gemüt*, and thus involves all the faculties in the first sense. For that reason, all five faculties<sub>1</sub> are represented under each faculty<sub>2</sub>. Sensibility and Imagination are separated from Understanding, Judgment and Reason by a line because they differ from the latter in being intuitive rather than discursive. Kant expresses this by calling them the “lower” faculties. Sensibility is distinct from Imagination for being receptive rather than spontaneous. The three higher faculties are each individuated by the distinct logical and epistemic functions for which they are considered to be responsible. According to Kant, the understanding is the faculty of concepts or of rules. I suggest we understand this, in line with the discussion of the three mental operations above, as the faculty for grasping concepts. To be sure, a concept is only grasped if it is already relatable to other concepts. Thus, the understanding is the capacity to deal with concepts alone. The power of judgment is distinct not because it can relate two terms, in *casu* concepts, but because it can subsume something under a concept. Judgment is thus the capacity to subsume particulars under concepts. In the next chapter, we will see more implications of this idea. For now, it is important to see that, since Kant believes (as I argued in 6.3.1) the understanding to be incapable of grasping particulars, another faculty is necessary to relate concepts and particulars. Finally, reason is not just the capacity for syllogism, it is the capacity for systematicity (as we saw in 5.3.1). Kant believes that no concept or judgment of itself need be considered as located in a general system of concepts. Nevertheless, he believed that it is a logical requirement that all concepts are thoroughly connected. This idea is represented in the idea of the absolute or that of systematicity, which belong to reason alone. Because of this capacity, when we subsume an object under a concept, we *ipso facto* assign it a place in a whole system of concepts, i.e. subsume it, indirectly, under the system of concepts.

In 6.3.1, we saw that Kant is unlikely to regard the lower and the higher faculties as reducible to each other or to a common ground. Moreover, that Kant insists on the

spontaneity of the imagination, and at times calls it a fundamental power, suggests that this faculty is not reducible to sensibility and understanding. Finally, Kant seems to believe that the three higher faculties are not unifiable either. The first reason is the distinctness of their objects and operations, and the second is the fact that the faculties<sub>2</sub> are not unifiable. Kant claims that the faculties<sub>2</sub> are distinct because, in each of them, a different higher faculty is legislative for the whole of the soul or *Gemüt* (AA V: 196-197). In my table above, this is represented by the fact that, under each faculty<sub>2</sub>, a different higher faculty<sub>1</sub> is set in boldface. Thus, the understanding is legislative for the faculty of cognition, i.e. for the theoretical usage of reason, the power of judgment is legislative for the feeling of pleasure and displeasure, and reason is legislative for the faculty of desire, i.e. reason in its practical usage. The three faculties<sub>2</sub> could thus not be distinct if their governing faculties<sub>1</sub> were not distinct. As a result, the mutual distinctness of the three higher faculties follows from the mutual distinctness of the three faculties in the sense of the quote from the unpublished introduction to the *Critique of the Power of Judgment*.

All of this suggests that, although Kant believed that there is definite benefit in reducing various epistemic functions to a limited number of general functions, he did not believe we could in fact reduce them to a single fundamental power, as Wolff would have liked. Considered from a logical or epistemological point of view, the soul consists of several qualitatively distinct faculties which have to work in concert to achieve their goal. For each general goal that the “soul” in this sense can set itself, one faculty is assigned a position of leadership among equals, and its claims are constrained by the fact that different faculties have this leadership in different engagements. This kind of unity, distinct as it is from the one attributed to the soul by Kant’s rationalist predecessors, is one that we must now try to characterize.

## 6.4 The System of Faculties as an Organic Unity

In the preface to the second edition of the *Critique of Pure Reason*, Kant makes a thought-provoking comment on the nature of the soul as it is considered in the critical philosophy:

pure speculative reason is, in respect of principles of cognition, a unity entirely separate and subsisting for itself, in which, as in an organized body, every part exists for the sake of all the others as all the others exist for its sake, and no principle can be taken with certainty in one relation unless it has at the same time been investigated in its thoroughgoing relation to the entire use of pure reason. (CPR B xxiii)

It is thought-provoking because it gives us a clue as to the kind of unity exhibited by the soul, namely as an organic unity. Two objections might immediately be raised against taking this passage in such a manner. The first is that it does not speak about the unity of the soul, but rather of the unity of speculative reason. This is merely a terminological confusion, since Kant often uses the term “speculative reason” to refer to the soul in its theoretical use, in the sense of what the *Critique of the Power of Judgment* called the “faculty of cognition”, and he often uses “reason” as synonymous for the whole of what rationalists prefer to call the soul, at least when he considers it from an epistemological point of view.

The other objection is that of the scope of the metaphor. Here, Kant aims to show that we can arrive at a complete and non-redundant list of all the principles of knowledge because these principles are thoroughly interrelated. The feature of an organized body that he is capitalizing on is its functional coherence. If we adhere to a functional perspective on an organism, as Kant does, then the organism consists of a number of functions, each of which contributes to the whole, and none of which are redundant. Moreover, taken together, these various functions suffice for the functioning of the organism. Kant thus seems to claim, first of all, that the principles of cognition are complete and non-redundant just like the various functions of a functioning organism are.

But Kant’s comment goes further than this. He not only states that we can arrive at a complete list of such principles in this way, he also warns us that this means that we cannot consider any of the functions in isolation. This claim is tantamount to Kant’s famous architectonic demand, that of the systematic coherence of his philosophy. Here, this architectonic unity is clearly understood in terms of an *organic* unity. We may wonder what further implications this has for his theory.

The employment of the organic metaphor in the context of the architectonic provides us with a further analogy between an organism and the system of faculties and principles, namely that of the specificity arising in it. As we saw in chapter 4, the different functions of an organism do not solely come together in a coordinated whole – they specify and produce each other. Taken in isolation, none of the functional parts of an organism would be functional, or would even have or maintain the specific structure it requires in order to function in the way they do. I therefore suggest that we understand the organic coherence of speculative reason also as one of reciprocal production and specification of the many functions. At least in one context does this become apparent, namely in that of the schematization of the categories. Without sensibility and imagination, the categories would never acquire their specific schematized forms, and would remain empty and general. I believe many such cases of interspecification can be given, and that the table of categories is probably better understood in this manner.

The idea of interspecification indicates a certain departure of Kant's from Crusius' faculty pluralism. As we saw in 6.2.2, Crusius conceived of the interrelation of the many powers of the soul on analogy with the interrelation of powers in dynamics. This means that vivification and composition are the primary mechanisms leading to the complexity of the system of cognition. When Kant departs from the physical and dynamic language in favour of the organic, this may indicate a fundamental change in the conception of the interrelation and interanimation of mental powers. In chapter 4, we saw that Kant believed the complexity and reciprocal causation at work in an organism to be principally irreducible to the interrelation of mechanic and dynamic forces. This is because the latter already enter into the complex fully specified, and simply compose their various specificities. Organic parts and powers, instead, acquire new specificity in and are even produced and reproduced by the organic whole in which they figure. This suggests that, where Crusius was still content to speak of faculties that are specifiable in isolation, Kant regards the faculties as conceivable only within the whole of the *Gemüt*.

We can find even further interesting consequences of the analogy between an organism and the system of faculties when we consider that Kant explicitly calls the faculty of reason an "organ" (*CPR* B xxxvii). He does so in the context of his criticism of Christian Wolff, whom he charges with neglecting to engage in a critique of pure reason itself. Now of course, the word "organ" might be meant in the general sense of "instrument", but, taken together with the already quoted passage from the same preface, it is likely to mean more. In fact, taking it more literally might even add to our understanding of Kant's criticism of Wolff. I suggest that Kant may be criticizing Wolff for not understanding that reason is but one organ amongst many. If this is true, the problem is that Wolff attributed to reason capacities that it does not itself have, such as that of presenting us with particulars without the mediation of other faculties, and of conflating the various distinct epistemic functions of the mind. This criticism would be understandable in the light of Kant's refusal of the Wolffian theory of an absolutely fundamental power.

This relation between the organic metaphor and the doctrine of faculties has yet another benefit. In the last subsection, I argued that Kant preferred to think of the mind or the soul, in so far as epistemology and logic are concerned, as a unity that is different from the unity of a simple substance. The Wolffian theory was that, in order to act as a unity, the faculties must themselves reduce to a single faculty, to a single governing and individuating principle, namely the representative power. Kant seems to envisage a unity of coordination and cooperation that is not grounded in such a grounding principle. In chapter 4, we saw that the unity of a natural purpose unity should not be thought of as apart from, or over and above, the multiplicity of parts of which it consists. It is rather a teleological, i.e. functional, i.e. normative unity of the diversity. Kant's picture of the system of faculties seems to have a similar structure. In 6.3.2, I



argued that Kant wished to dissociate the idea of the unity of thought from that of the simplicity and substantial unity of that which does the thinking. This is an understandable move if he was indeed trying to reconceive the unity of thought as an organic unity, a functional unity of diverse, interspecifying parts.

This raises an important question, however. In the case of an organism, its unity was considered normative because it could *fail* to obtain. That is to say, an organism can be diseased, hurt and even killed. I argued that the normative unity of an organism should not just be understood as the wonderful adaptedness of its parts, but also as the ability to overcome some distortions of the general functionality, and even to alter its normative unity to deal with recalcitrant situations. In other words, I have argued that it is essential to an organism that it is epigenetic, i.e. productive and adaptive. We should therefore inquire whether Kant also thought of the system of faculties in this latter way, or whether he instead thought that the soul, being immaterial and transcendent, is immune to the vicissitudes that organized bodies are constantly prompted to overcome. In other words, we must inquire whether the unity of the system of faculties can fail to obtain, or is rather eternally and perfectly adapted to its task: that of knowing. This question will be dealt with in the next and last chapter of this dissertation.

## 6.5 Conclusion

In this chapter, I have argued that Kant sought to defend, against Christian Wolff and his followers, the idea that we need to accept a multiplicity of faculties that are not reducible to each other. In doing so, Kant sided with Crusius, although he also heeded Tetens' attempt to maximize the unity of the faculties. In doing so, he ended up with five distinct faculties, two that are intuitive and three that are discursive. I have provided reasons for the irreducibility of these faculties, and have specifically discussed the distinction between the discursive and the intuitive at great length. In that context, I have argued that the specific and irreducible contribution of sensibility to knowledge is due to its intuitive nature, and not its receptive nature. The upshot is that the major opposition in Kant's epistemology is not that between the given and the thought, but rather that between the particular and the general. I have also argued that the unity of the mind in Kant's critical philosophy is best understood as an organic unity. The further implications of these theses will be developed in the next chapter.



**Part III:**

**Systematicity and its Historicity**



## Chapter 7 The Epigenesis of the Faculties

*Kant beschränkt sich mit Vorsatz in einen gewissen Kreis  
und deutet ironisch darüber hinaus.*  
- Goethe

In this chapter, I will answer the two questions with which I closed part 2 by showing how the *Critique of the Power of Judgment* elaborates on the organic conception of the *Gemüt*. Before doing so, I offer, in 7.1, a brief discussion of some major figures and trends in 18<sup>th</sup> century aesthetics. This discussion is meant to clarify the goal of the “Critique of the aesthetic power of judgment”, where Kant offers his theory of the harmony of the faculties. It is to this latter theory that I devote 7.2. In 7.2.1, I discuss reflective judgment, and show that it is primarily concerned with the relation between the faculties. In 7.2.3, I then introduce the notion of the harmony of the faculties, and argue that we need to read it as the purposive unity of the *Gemüt*. In 7.2.3, I argue that this purposive unity is contingent, and that it is strived after by the *Gemüt*. This answers the first question of part 2: the system of faculties is indeed also like an organism because its normative unity is contingent. In 7.3, I then answer the second question, arguing that the system of faculties is indeed epigenetic because the sublime opens the system up to history and change. In 7.4, I then try to bring together the main lines of dissertation and offer a picture of Kant’s epigenesis and historicity of the a priori.

### 7.1 Aesthetics and Cognition in the Eighteenth Century

It is one of the lasting commonplaces of the historiography of philosophy that the discipline of aesthetics emerged in the eighteenth century. Even though it is ludicrous to say that there was no philosophical reflection on the nature of beauty or the experience thereof before the age of Enlightenment, it is arguable that this reflection acquired not one, but many new impetus during that time. In two distinct traditions, the

issue of the aesthetic came to be a central concern, namely in the Scottish Enlightenment and in German *Schulphilosophie*. In Kant, we see a confluence of these two influences, specifically in his attempt to answer the question as to the nature and source of aesthetic experience. In this section, I will provide a brief overview of some major positions in the debate that are likely to have been in Kant's mind. I will present them with a specific issue in mind, namely that of whether beauty is experienced through the senses or through cognition.

On this issue, it is tempting to immediately suggest that, in general, British authors inspired by John Locke took aesthetic experience to be an experience of the senses, whereas German authors working in the tradition of Leibnizian rationalism had an intellectualist perspective on beauty. This suggestion may aid us in understanding how Kant's solution to the problem is an application of his more general tendency to overcome the empiricist-rationalist divide. It does so, however, at the expense of the complexity of the debate. As we will see, although some positions can easily be read through this stark opposition, the more interesting ones that Kant engaged with were more nuanced. In this section, I will therefore present both the extreme and the more complex position of both traditions.

Leibniz has often been regarded as the major source for the intellectualist strand in German aesthetics, given his belief "that fundamentally pleasure is a sense of perfection, and pain a sense of imperfection, each being notable enough for one to become aware of it." (G V: 180; Leibniz 1996: 194) Pleasure can be gained from the perception or awareness of perfection in an object, as is the case in music:

Music charms us, even though its beauty consists only in the harmonies of numbers and in a calculation that we are not aware of, but which the soul nevertheless carries out, a calculation concerning the beats or vibrations of sounding bodies, which are encountered at certain intervals. The pleasures that sight finds in proportions are of the same nature, and those caused by the other senses amount to something similar, even though we might not be able to explain it so distinctly. (G VI: 605-606; Leibniz 1989: 212)

Leibniz suggests that aesthetic pleasure is an awareness of a feature that can be cognized by the intellect, namely the perfection and harmony of the perceived structure or entity, but that this awareness needn't include awareness of the precise nature of this perfection. In the language of Leibniz's version of the theory of ideas, the pleasure we take in perfection requires that our perception of this perfection be clear, but not that it be distinct. Thus, the pleasure in the perception of beauty is analogous to a secondary quality. According to Leibniz, secondary qualities are confused ideas of the primary qualities with which they correspond, and if we were to fully analyze, i.e. make distinct, that perception, we would see *how* the secondary qualities are due to the primary ones. Similarly, if we were to fully analyze our pleasure in the beauty of an

object, we would arrive at the distinct perception of the harmony and perfection of that object.

The same opinion is to be found in Christian Wolff's writings, for instance in the *German Metaphysics*, where Wolff defines pleasure (*Lust*) as an *intuitive knowledge of perfection* (Wolff 1747: 247). All three terms indicate the specificity of the Leibnizian-Wolffian perspective on the aesthetic experience. First of all, pleasure is a form of knowledge, and therefore cognitive. Moreover, it is a cognition of a feature of the perceived object, namely its perfection. The latter notion is central to Wolff's philosophy, and he defines it as the harmony of the diverse (Wolff 1747: 78). The pleasure we take in perceiving such an object thus consists in the cognition of the harmony or organization of the object. Of course, Wolff admits that we can take pleasure in objects that fail to have such perfection, because the mere illusion of perfection can constitute pleasure as well (Wolff 1747: 248). The cognitive nature of aesthetic pleasure has two important consequences in Wolff's account. First of all, it means that our appreciation is not only proportionate to the degree of perception of the object, but also to our knowledge of that perception (Wolff 1747: 250). Secondly, it means that such pleasure occurs as well in the cognitive pursuits of science and discovery (Wolff 1747: 251-252).

Wolff specifies, however, that pleasure is a specific kind of cognition of the perfection of an object, namely an intuitive (*anschauende*) one. As we saw in 6.2.1, he believed that ultimately there is only one fundamental mental power, namely the representative force, and that ideas differ only in their degree of clarity and distinctness. Intuitions are therefore ideas just like concepts, but they are confused ideas. This is why Wolff argues that pleasure does not require that we have distinct knowledge of the object, and only that we have clear knowledge of it – it suffices that we have an intuition of the perfection (Wolff 1747: 252). It is also why Wolff believes we can have pleasure in the absence of perfection, namely because we may be confused. His opinion is then that our pleasure in the perfection of things becomes more adequate to the extent that our knowledge of those things becomes more distinct (Wolff 1747: 254-255).

The Leibnizian-Wolffian view is thus that aesthetic pleasure is a cognitive state: it is an awareness of a feature of the object. Our sense of pleasure corresponds with an objective feature of the object, namely its degree of perfection, and is moreover a recognition of that feature.<sup>1</sup> We can have both intellectual and sensible knowledge of

---

<sup>1</sup> Some qualification is required with regard to the status of the concept of beauty. Wolff defines beauty as “the perfection of a thing, insofar [quatenus] as with this force it is apt to produce pleasure in us” (Wolff 1738: 420; my translation). The second part of the sentence implies that Wolff does not directly identify beauty with perfection, but rather with an aptitude of perfection to produce some things in us. It may seem as if he thereby makes beauty into a dispositional or a relational property. This may be the case, but it remains to be

that feature, but the more intellectual this knowledge is, the higher the chances that it is also adequate. The adequacy of an aesthetic feeling is the degree to which it corresponds with the feature of the object rather than with a confusion regarding whether that object has the feature. This strikes us as an intellectualist theory of aesthetic experience, since it degrades sensibility as a lower, less trustworthy form of cognition, and stresses the similarity between the pleasure we take in a work of art and the one we take in the elegance of a mathematical theorem or scientific theory.

It is common to note that the highly rationalist and cognitivist nature of German aesthetics was somewhat mitigated by Alexander Gottlieb Baumgarten, as Frederick Beiser does:

Within the rationalist tradition, Baumgarten occupies the middle ground. His core values are fundamentally rationalist. His ideal of knowledge is entirely intellectual: we know best what we clearly and distinctly conceive, or what we can demonstrate through reason alone. Reason always remains for him the *higher* faculty of cognition, the senses the *lower* faculty. Nevertheless, Baumgarten gives a much greater value than Wolff or Gottsched to the distinctive qualities of sensible cognition. (Beiser 2009: 122)

This claim is made because, although Baumgarten adopts many typically Wolffian standpoints, he does depart from his great example in one major aspect, namely that of the status of sensibility. When Baumgarten coined the idea of a separate science of aesthetics, he defined it as “the science of sensible knowledge” (Baumgarten 1750: 1) What is considered revolutionary about this is that Baumgarten felt the need to defend a separate science of knowledge gathered through the lower faculties, whereas it is considered a core tenet of rationalism that only the higher faculties can be trusted with the acquisition of knowledge. In Leibnizo-Wolffianism, this is a *fortiori* the case, because sensibility is considered to deliver ideas that are not different in kind from those of the higher faculties, and only differ from them in their degree of distinctness. If this is true, there does not seem to be any specific benefit to be reaped from pursuing the specific kinds of knowledge delivered by sensibility. Thus, although Baumgarten ultimately agrees on the lower status of the lower faculties and the confused nature of the representations of sensibility (Baumgarten 1779: 180-181), and remains a rationalist in the sense of the present debate, he did seek to mitigate this rationalism. What remains

---

noted that in Leibnizian and Wolffian metaphysics, the latter kinds of properties are strictly grounded in non-dispositional and non-relational properties such as perfection. My conjecture is that beauty is not real, but rather a well-founded phenomenon – well-founded because it is founded in the perfection of the thing. Indeed, the “*quatenus*” formula is a typical Leibnizian tool of expressing relational properties as grounded in non-relational properties (cf. Parkinson 1965: 49).



unchanged, however, is the *cognitivist* aspect of his aesthetics: beauty is still the intuitive cognition of the perfection of the object.

Like the tradition of 18<sup>th</sup> century German aesthetics, British, and most certainly Scottish Enlightenment thought on the topic is notoriously rich, and can therefore not be treated here in full. I will only briefly discuss the contributions of two figures, namely Francis Hutcheson and Edmund Burke, because they have visibly informed Kant's reflections on the subject, as we will see. A comparison between their differing views will yield a better picture of just how difficult it is to speak of a straightforwardly empiricist theory of aesthetics. Above, I have described the German aesthetic tradition of the *Schulphilosophie* as deeply rationalist and cognitivist. By this I meant that it preferred reason as a faculty of knowledge, and that it regarded the aesthetic experience as a cognition of a feature of an object. I will argue here that Hutcheson and Burke are both sensualist/empiricist and non-cognitivist in their aesthetics, despite their differing conceptions on the nature of beauty.

Francis Hutcheson is perhaps best known for his suggestion that Taste is not only a sense, rather than an intellectual faculty such as judgment, but that it is actually an *internal* sense. He gives two reasons for contrasting this sense with the external senses. The first reason is that having a particular external sense does not automatically make one fit to appreciate the associated arts (Hutcheson 2004: 23). The mere capacity to hear, for instance, does not make one apt to properly appreciate the harmonies of music. In making this comment, Hutcheson suggests that the appreciation of beauty does not simply accompany our basic senses. The second reason for contrasting inner and outer sense is that the feeling of beauty can occur in situations where the senses are not really concerned, like in the sciences (Hutcheson 2004: 24).

Nevertheless, Hutcheson insists that the sense of beauty is a sense rather than an intellectual judgment. The first reason for this is that

the Ideas of Beauty and Harmony, like other sensible Ideas, are necessarily pleasant to us, as well as immediately so; neither can any Resolution of our own, nor any Prospect of Advantage or Disadvantage, vary the Beauty or Deformity of an Object. (Hutcheson 2004: 25)

This conclusion follows from Hutcheson's opening premise that

Those Ideas which are rais'd in the Mind upon the presence of external Objects, and their acting upon our Bodys, are call'd Sensations. We find that the Mind in such Cases is passive, and has not Power directly to prevent the Perception or Idea, or to vary it at its Reception, as long as we continue our Bodys in a state fit to be acted upon by the external Object. (Hutcheson 2004: 19)

The first reason for the sensible nature of taste is, then, that in an aesthetic experience, our feeling of beauty does not depend on our free will, and is an involuntary response to

the perceived object rather than a feeling that we can conjure in ourselves at will. In this, taste most closely resembles our sensations in being *passive* and *receptive* rather than *active* and *spontaneous*.

The second reason why Hutcheson believes taste to be a sense rather than an intellectual faculty is that it is *immediate* rather than *mediated by knowledge or concepts*:

This superior Power of Perception is justly called a Sense, because of its Affinity to the other Senses in this, that the Pleasure does not arise from any Knowledge of Principles, Proportions, Causes, or of the Usefulness of the Object; but strikes us at first with the Idea of Beauty: nor does the most accurate Knowledge increase this Pleasure of Beauty, however it may super-add a distinct rational Pleasure from prospects of Advantage, or from the Increase of Knowledge. (Hutcheson 2004: 25)

Hutcheson thus argues that the pleasure we have in the beauty of a thing is strictly distinct from the theoretical or practical interests we may have in it, although it may accompany the specific pleasures connected to those latter interests. This is an important comment because it urges for the autonomy of the aesthetic from practice and theory alike – important because, but not just because, it would also become a tenet of Kant's aesthetics in the *Critique of the Power of Judgment*.

But there is another aspect to the distinction from rational interest that reveals a clear contrast with Wolff's position in the *Deutsche Metaphysik*. As we saw above, Wolff believed that the degree of aesthetic appreciation does not just vary with its degree of perfection, but also with the distinctness of our knowledge of this perfection. He thus believed that our pleasure in the beautiful is increased by our knowledge. According to Hutcheson, the pleasure of beauty cannot be increased in this way except perhaps indirectly. Our sense of beauty is an involuntary and immediate response, and it cannot be changed by our knowledge. It can, however, be counteracted by it in various ways. This is why Hutcheson argues in a later section that Education and Custom can do fairly little for our sense of taste, except perhaps by removing the prejudice which distorts our appreciation. This too marks a distinct difference with Wolff, who believed that we can be mistaken about the beauty of an object by the confusion of our idea of that object. According to Hutcheson, such error is strictly speaking impossible, and our errors in aesthetic judgment are entirely due to the fact that we are wrongly tempted to associate another idea, which bears on our judgment of value, with the idea of the object proper (Hutcheson 2004: 67). Thus, it is not the nature of our idea of the object that leads us into error, but rather the web of ideas and prejudice within which it soon finds itself embedded, and our reasonings on the object's pertinence to other objects and aims.

If we can label Edmund Burke a sensualist in aesthetics, it is certainly not on the ground that he believed taste to be a sense, as Hutcheson obviously did. In fact, Burke believed taste did not pertain to one faculty alone:

On the whole it appears to me, that what is called Taste, in its most general acceptation, is not a simple idea, but is partly made up of a perception of the primary pleasures of sense, of the secondary pleasures of the imagination, and of the conclusions of the reasoning faculty, concerning the various relations of these, and concerning the human passions, manners and actions. All this is requisite to form Taste, and the ground-work of all these is the same in the human mind; for as the senses are the great originals of all our ideas, and consequently of all our pleasures, if they are not uncertain and arbitrary, the whole of Taste is common to all, and therefore there is sufficient foundation for a conclusive reasoning on these matters. (Burke 1764: 30-31)

However, this passage already contains important hints about the general empiricist and sensualist bend of Burke's thought. First of all, there is the idea that taste has universal principles because the senses of different men are usually similarly constituted. This means that Burke believes our senses to be functionally identical or sufficiently similar as they operate normally, and that they diverge only in cases of abnormality such as disease or injury. Such stability is not to be found in reasoning, Burke believes, and is specific to the senses (Burke 1764: 33-34). This leads to the good empiricist maxim that men agree insofar as they are sensible, and disagree insofar as they are reasoners, and contradicts the rationalist assumption, which we saw at work in Wolff's aesthetics, that men disagree insofar as they are sensible, and agree insofar as they are reasoners.

The second indicator of Burke's empiricism is the related idea that the pleasure of beauty comes from the lower, and not the higher faculties. Indeed, Burke states in the quoted passage that taste is concerned with the primary pleasures of the senses and the secondary senses of the imagination. The former are the pleasures brought about in us by the perception of certain objects, and the latter are brought about by our noticing of similarities between objects. It is for this reason that imagination is more properly and commonly a source of pleasure than judgment:

But in the imagination, besides the pain or pleasure arising from the properties of the natural object, a pleasure is perceived from the resemblance, which the imitation has to the original; the imagination, I conceive, can have no pleasure but what results from one or other of these causes. And these causes operate pretty uniformly upon all men, because they operate by principles of nature, and which are not derived from any particular habits or advantages. Mr. Locke very justly and finely observes of wit, that it is chiefly conversant in tracing resemblances; he remarks at the same time; that the business of judgment is rather in finding differences. It may perhaps appear, on this supposition, that there is no material distinction between the wit and judgment, as they both seem to result from different operations of the same faculty of *comparing*. But in reality, whether they are or are not dependent on the same power of the mind, they differ so very materially in many respects, that a perfect union of wit and judgment is one of the

rarest things in the world. When two distinct objects are unlike to each other, it is only what we expect; things are in their common way; and therefore they make no impression on the imagination: but when two distinct objects have a resemblance, we are struck, we attend to them, and we are pleased. The mind of man naturally has a far greater alacrity and satisfaction in tracing resemblances than in searching for differences; because by making resemblances we produce *new images*, we unite, we create, we enlarge our stock; but in making distinctions we offer no food at all to the imagination; the talk itself is more severe and irksome, and what pleasure we derive from it is something of a negative and indirect nature. (Burke 1764: 17-19)

Here, Burke argues for the real distinctness of the faculties of imagination and understanding by indicating their different effects in aesthetics: the noticing of similarities, which is the proper business of imagination, is productive of pleasures, whereas the noticing of distinctions, judgment's function, does not contribute to the aesthetic feeling except in a negative way. Both Burke and Hutcheson therefore believe that, although the higher faculties can make contributions to taste, these contributions are mostly negative.

Burke and Hutcheson also agree in their non-cognitivism in aesthetics, despite their widely differing conceptions of beauty. In Hutcheson, we find a theory of beauty that is remarkably similar to the one we saw in Wolff and Baumgarten, namely the idea that beauty consists in what Wolff called "perfection":

what we call Beautiful in Objects, to speak in the Mathematical Style, seems to be in a compound Ratio of Uniformity and Variety: so that where the Uniformity of Bodys is equal, the Beauty is as the Variety; and where the Variety is equal, the Beauty is as the Uniformity. (Hutcheson 2004: 29)

What prompts in us the aesthetic appreciation of an object is its high degree of uniformity amid variety, of harmony, of proportion, at least insofar as it is considered in itself and not as an imitation of another object. Mind however that Hutcheson did not believe beauty to consist in perfection; he rather takes beauty to be the analogue of a secondary quality in the inner sense corresponding to the primary quality of uniformity amidst variety (Hutcheson 2004: 27). This already marks a major distinction with the rationalist tradition: according to Hutcheson, the experience of beauty does not consist in, and does not even require, the cognition or recognition of the uniformity amidst variety, however confused:

But in all these Instances of Beauty let it be observ'd, That the Pleasure is communicated to those who never reflected on this general Foundation; and that all here allerdg'd is this, "That the pleasant Sensation arises only from Objects, in which there is Uniformity amidst Variety:" We may have the Sensation without knowing what is the Occasion of it; as a Man's Taste may suggest Ideas of Sweets,

Acids, Bitters, tho he be ignorant of the Forms of the small Bodys, or their Motions, which excite these Perceptions in him. (Hutcheson 2004: 35)

The idea is that beauty is not experienced as a cognition, but rather that is simply a fact about our inner sense that it is prompted to produce the idea of beauty when it is confronted with an object displaying uniformity amidst variety. This connection between beauty and uniformity is not a priori or cognitive, it is a contingent fact decided by the author of our nature, God:

There seems to be no necessary Connection of our pleasing Ideas of Beauty with the Uniformity or Regularity of the Objects, from the Nature of things, antecedent to some Constitution of the Author of our Nature, which has made such Forms pleasant to us. Other Minds may be so fram'd as to receive no Pleasure from Uniformity; and we actually find that the same regular Forms seem not equally to please all the Animals known to us, as shall probably appear afterwards. (Hutcheson 2004: 46)

This marks a further difference with rationalism, which becomes clear if we relate this discussion with my treatment of the debate on antisystematicity in 5.2. There, we saw that rationalists tend to regard systematicity as an a priori virtue of the world, i.e regard uniformity amidst variety as a maxim by which God himself operates. On this assumption it is natural to suppose that we are automatically pleased when we recognize uniformity amidst variety, because the latter is inherently pleasing. According to empiricists, there is no a priori reason why God would have to abide by the rule of uniformity amidst variety, and would not be tempted to create a world that is inconceivable in its richness. If this is true, the desirability of uniformity amidst variety is decided by God, and not a ground of God's decisions. Hutcheson here embodies the empiricist tradition by noting that it is neither necessary that the world exhibit uniformity amidst variety, nor that we take pleasure in uniformity amidst variety. He does reveal himself to be profoundly theistic by suggesting that the fact that we, who are contingently appreciate uniformity amidst variety, find ourselves in a world, or a stretch of the world, which greatly exhibits uniformity amidst variety, implies that a wise intentional artificer has created both our natures and the world in which we live with an eye towards their mutual conformity (Hutcheson 2004: 46-47). This idea manifests itself even more clearly in his discussion of the "final causes" of inner sense, by which he means the divine intentions for inner sense<sup>2</sup>:

---

<sup>2</sup> Remember from 4.1 that these two statements are only synonymous in a very specific programme, which we can associate with that of the Boyle lectures and the general tendency to interpret teleology as design, and design as a mechanical structure intended and disposed by God.

That supposing the Deity so kind as to connect sensible Pleasure with certain Actions or Contemplations, beside the rational Advantage perceivable in them; there is a great moral Necessity, from his Goodness, that the internal Sense of Men should be constituted as it is at present, so as to make Uniformity amidst Variety the Occasion of Pleasure.” For were it not so, but on the contrary, if irregular Objects, particular Truths, and Operations pleased us, beside the endless Toil this would involve us in, there must arise a perpetual Dissatisfaction in all rational Agents with themselves; since Reason and Interest would lead us to simple general Causes, while a contrary Sense of Beauty would make us disapprove them: Universal Theorems would appear to our Understanding the best Means of increasing our Knowledge of what might be useful; while a contrary Sense would set us on the search after particular Truths: Thought and Reflection would recommend Objects with Uniformity amidst Variety, and yet this perverse Instinct would involve us in Labyrinths of Confusion and Dissimilitude. And hence we see “how suitable it is to the sagacious Bounty which we suppose in the Deity, to constitute our internal Senses in the manner in which they are; by which Pleasure is join’d to the Contemplation of those Objects which a finite Mind can best imprint and retain the Ideas of with the least Distraction; to those Actions which are most efficacious, and fruitful in useful Effects; and to those Theorems which most enlarge our Minds.” (Hutcheson 2004: 80-81)

Hutcheson argues here that, were we not naturally inclined to seek unity and similarity, and take pleasure in disunity and variety alone, we would constantly be dissatisfied with our rational labours. The theism here consists in the idea that God has been merciful in making our aesthetic desires match our rational natures and the rational structure of the world. In other words, God has disposed our inner sense thus that, although it does not consist in cognition, it best matches our cognitive needs.

It is this last connection with cognitivism in aesthetics that Burke attacks when he proposes his own theory of the beautiful and the sublime. When reading Burke’s discussion of what beauty is *not*, it is hard not to think that it is a direct criticism of Hutcheson, since he argues there that beauty consists neither in proportion nor in perfection (Burke 1764: 202-204). Instead, Burke believes “that beauty is, for the greater part, some quality in bodies, acting mechanically upon the human mind by the intervention of the senses” (Burke 1764: 210). At first, this may seem not to differ that greatly from Hutcheson’s account, since the latter too implied that beauty is a secondary quality produced by some primary qualities. But for Hutcheson, this connection was one between a clearly delineated property of the object and an idea of a specific internal sense. Burke’s account is much more physiologicistic, as he is not content to merely state that God has installed a contingent connection between a property and an idea. What Burke wants to understand is how our physiology and anatomy are responsible for our taking aesthetic pleasure in certain things rather than others. The difference in view becomes all the more clear in Burke’s discussion of the sublime.

There, the sublime is characterized as connected to pain rather than pleasure, and pain and pleasure as connected to the sense of self-preservation in man (Burke 1764: 160). The specifically aesthetic nature of the feeling of the sublime is that we take delight in the feelings of pain and terror prompted in us by the dangerous in absence of real danger (Burke 1764: 84-85). This is important because it allows Burke both to explain aesthetic feelings as parasitic upon physiological effects and biological responses, and to stress that they are not themselves biologically interested in the way that “more base” pleasures and delights are.

I have provided this overview in order to indicate some major issues which inform Kant’s discussion of the aesthetic and the harmony of the faculties, issues like the questions whether taste is better understood as intellectual or as sensual, cognitive or non-cognitive, as connected to objective features of the world or not, as a physiological or a specifically aesthetic response, etc... In the rest of this chapter, we will see Kant’s peculiar way of responding to these various questions, and how he thereby lays the basis for understanding the system of faculties as an organic whole that is genuinely normative and epigenetic.<sup>3</sup>

## 7.2 The Functional Unity of the Faculties

### 7.2.1 Reflective Judgment

If we want to understand the specificity of Kant’s position in the *Critique of the Power of Judgment*, we need to gain a better understanding of the faculty of which it is the *Critique*. As I showed in 6.3.3, Kant believed that there were ultimately three distinct global orientations of the *Gemüt*, which he called the faculty of cognition, the faculty of desire and the feeling of pleasure and pain, and that in each of these faculties one of the three higher faculties was dominant. The *Critique of the Power of Judgment* is devoted to the feeling of pleasure and pain, in which the power of judgment is dominant. Hence,

---

<sup>3</sup> My discussion here has thus been motivated by my general concern with the nature of the system of faculties, rather than with philosophical aesthetics. This is the reason for my omission of a historical setting that is essential to the proper understanding of Kant’s contributions to the latter, namely the debate on the standard of taste. In these matters, David Hume and Henry Home, Lord Kames, deserve extensive treatment. Since I will not be focusing on the issues of the *sensus communis* and the deduction of taste, I have omitted such discussions, and focused on the relation between aesthetics and epistemology and with teleology.

the third *Critique* is an inquiry into the competences and the a priori principles of judgment.

Kant had already introduced the idea of a specific faculty of judgment in the first *Critique*, although there he still understandably considered it from the perspective of its specific contribution to the theoretical use of reason under the guidance of the understanding. As a result, the doctrine of the power of judgment there was merely a discussion of the principles of the understanding (*CPR* A 132 / B 171). Nevertheless, he did contrast understanding and judgment there:

If the understanding in general is explained as the faculty of rules, then the power of judgment is the faculty of **subsuming** under rules, i.e. of determining whether something stands under a given rule (*casus legis*) or not. General logic contains no precepts at all for the power of judgment, and moreover cannot contain them. For **since it abstracts from all content of cognition**, nothing remains to it but the business of analytically dividing the mere form of cognition into concepts, judgments and inferences, and thereby achieving formal rules for all use of the understanding. Now if it wanted to show generally how one ought to subsume under those rules, i.e. distinguish whether something stands under them or not, this could not happen except once again through a rule. But just as this is a rule, it would demand another instruction for the power of judgment, and so it becomes clear that although the understanding is certainly capable of being instructed and equipped through rules, the power of judgment is a special talent that cannot be taught but only practiced. (*CPR* A 132-133 / B 171-172)

Kant thus distinguishes judgment and understanding in that the former is the capacity to reason through and with rules, the latter is the capacity to subsume a given case under a rule. We should probably understand this as follows: logic allows us to determine general laws that govern the relations between concepts, provided that these relations are analytic. But the capacity to subsume a given case under a rule cannot be grasped under rules. The argument he gives here is that this would constitute a regress, for if there were rules for the application of rules, then there would have to be rules for the application of the rules for the application of rules.

It is difficult to see how we should interpret this argument to make it effective. Is Kant saying that the problem is that of the application of logical rules alone? In that case, the problem might be the following: we know that a subordinate concept contains its superordinate concept; what we do not know is whether two given concepts actually relate in this way. The issue with judgment then seems to be that, although the relations between concepts are governed by strict rules, we can never establish a rule to see in what relation two concepts stand. I find this argument unconvincing, because it is supposed that logic regards concepts insofar as they exhibit marks. If we therefore want to ask whether a given concept contains a given mark, then we are in a way asking whether that concept is that concept. Given Kant's conception of analyticity, it is



doubtful whether the question: “is that bachelor unmarried” requires a distinct contribution of judgment just to grasp that the concept of a bachelor contains that of being unmarried and that therefore we are dealing with a case of analytic truth.

I think it is better to understand the problem here through the definition of the power of judgment from the *Critique of the Power of Judgment*, namely as the power of subsuming the particular under the universal (AA V: 179). I suggest we read this in light of the discussion of 6.3.1, and see it as the power to subsume the particulars insofar as they are presented through the lower faculties to the universals, i.e. the concepts of the understanding. This allows us to make better sense of why subsumption is so problematic: we can easily relate two concepts because we can compare them in terms of the marks they contain. But we cannot compare a particular and a universal in this way because particulars are thought through intuitions, which are not concepts. This also appears to from Kant’s discussion of “the logical use of the understanding in general” in the *Critique of Pure Reason*:

Now the understanding can make no other use of [...] concepts than that of judging by means of them. Since no representation pertains to the object immediately except intuition alone, a concept is thus never immediately related to an object, but is always related to some other representation of it whether that be an intuition or itself already a concept). Judgment is therefore the mediate cognition of an object, hence the representation of representation of it. In every judgment there is a concept that holds of many, and that among this many also comprehends a given representation, which is then related immediately to the object. (CPR A 68 / B 93)

Kant is saying here that judgments always contain concepts, and that concepts never relate directly to objects, but that only intuitions do. This implies that there are two kinds of judgments: a judgment in which a concept is related to a concept, and one in which it is related to an intuition. In the first case, there is no real problem of subsumption, since all that is at stake is whether the two concepts relate, a question we can solve by analyzing the concepts. In the second case, however, there is a problem, since the relation between a concept and an intuition is one between two heterogeneous principles. An object can thus be subsumed under a concept if a property is intuited in it that is also thought of in the concept. The problem is that of relating “what is intuited in it” with “what is thought of it”, given that these two aspects are not clearly related.

There is a passage which seems to contradict this interpretation at first sight, namely the one from the opening of the *Schematism* chapter in the *Critique of the Power of Judgment*: “In all subsumptions of an object under a concept the representations of the former must be **homogeneous** with the latter, i.e. the concept must contain that which is represented in the object that is to be subsumed under it, for that is just what is

meant by the expression “an object is contained under a concept” (CPR A 137 / B 176). This seems to say that intuitions and concepts *can* be homogeneous after all. But this is based on a misunderstanding. As the rest of the chapter reveals, the homogeneity is not in fact between the intuition or object and the concept, but rather between the intuition or object and the schema of the concept. In fact, Kant argues that concepts and objects or images only ever relate through the mediation of a schema (CPR A 141 / B 180), which is a rule for determining the intuition in accordance with the concept. The reason why Kant believes intuitions and empirical concepts to be “homogeneous” is because he believes that the *schemata* of empirical concepts are homogeneous to intuitions. But much like there cannot be a rule for subsuming particulars and universals, so too we have no insight in the capacity to generate schemata.

All in all, I find it plausible that Kant thought of the power of judgment as the capacity to subsume intuitions under concepts by means of schemata, and that it is the difference in kind between intuitions and concepts that is the cause of the problem of subsumption. In the *Critique of the Power of Judgment*, however, Kant distinguishes between two different kinds of such subsumption of particulars under universals:

If the universal (the rule, the principle, the law) is given, then the power of judgment, which subsumes the particular under it (even when, as a transcendental power of judgment, it provides the conditions *a priori* in accordance with which alone anything can be subsumed under that universal), is **determining**. If, however, only the particular is given, for which the universal is to be found, then the power of judgment is merely **reflecting**.

The determining power of judgment under universal transcendental laws, given by the understanding, merely subsumes; the law is sketched out for it *a priori*, and it is therefore unnecessary for it to think of a law for itself in order to be able to subordinate the particular in nature to the universal. (AA V: 179)

The idea of a reflective judgment seems to be that of judging an object in absence of a concept for that object. Straightforward though this may sound, it raises serious interpretative difficulties when one dwells upon it. After all, it is a main tenet of the Kantian philosophy that we are never presented objects without concepts: each object-representation requires an intuitive and a conceptual element. It cannot be the case that in reflective judgment, we are considering a brute empirical given in search for a concept with which to grasp it, since we are never presented with such brute empirical givens. Or do we need to take this as evidence that Kant did succumb to the Myth of the Given when he introduced the idea of reflective judgment? A passage from the Dohna Wundlacken-Logic concerned with Kant’s third *Critique* doctrine of taste is explicit about this:

All our cognitions involve the following two things {two elements, one of which without the other yields no cognition}: 1. intuition (the interpretation" of the

concept, of thought); 2. concept. A pure concept like a pure rule does not yet yield any distinctness of cognition[;] this requires intuition, too. Conversely, intuition without concept is likewise nothing. For without [a concept] it would be as if it had seen nothing. (AA XXIV: 706)

As a result, we cannot simply claim that Kant forgot about this crucial piece of doctrine when he considered aesthetics.

An initial response to this is suggested by the fact that Kant connects the issue of reflective judgment with that of systematicity in the passages immediately following the one I just quoted. This initially suggests the following reading: in reflective judgment, we are considering not a brute empirical given, but rather an object as conceived under a highly specific empirical concept, and search for a more general concept under which to subsume it. The goal of reflective judgment is then that of unification: we attempt to find more general concepts through which to understand the particular in the universe, and try to unify our various specific concepts.

Plausible though such a reading might seem at first, it cannot be made to square with the general thrust of the third *Critique*, for later on, Kant speaks of reflective judgments in which only a single object is concerned. It does not seem fruitful to seek unificatory concepts and theories empirically by reflecting on a single instance or concept. Moreover, this interpretation makes the whole idea of aesthetic judgment impossible, as we will see further on.

In light of the above considerations, I will suggest the following reading: what we are trying to do in a reflective judgment is to come up with a concept that grasps an individual in its particularity, i.e. to think the particular (object) through the general (concept). This need not mean that we are presented with that object otherwise than through concepts: all that is required is that none of the various concepts with which we grasp it capture it adequately. In 5.4.1, I already suggested that we not read Kant's conception of the adequacy of a concept to an intuition or object as a correspondence between that concept and the representation. Instead, we may take a concept to be adequate if, in using it to grasp an object, our usage of that concept aligns with our uses of the other concepts through which we grasp that object or with other uses of the same concept in grasping other objects. On such a reading, we can be said to have no concept of an object if there is an incongruence between the various concepts through which we grasp it initially.

This makes the issue of reflective judgment not one of finding the general features of nature, but rather one of finding concepts that are ever better at grasping the specificity in nature. At first sight, this may conflict with the immediate context of the quoted passage, namely that of the concern with systematicity, which appears to be a concern with finding a coherent, unified picture of nature. In the light of 5.3.1, my response to this issue is to be expected, since I showed there that the *Critique of the Power of Judgment* attempts to find a conception of systematicity that allows us to square uniformity with

diversity. Moreover, in that section I characterized the principle of reflective judgment as a principle warranting analogical reasoning. In the present context, this means the following. Without the assumption of systematicity, the fact that none of the concepts through which we are grasping an object match up in applying them to this particular, suggests we need a different concept to grasp it. Unfortunately, we have no idea of where to start looking for such a concept, and what form it may take. With the assumption of systematicity, however, we may take the fact that these concepts at least seem adequate in some respects as a clue. That is to say we may assume that the object is at least *partially* grasped through these concepts, whereas in absence of the principle, we cannot be sure that these concepts are not *completely* inadequate. The assumption that the initial similarities can be used as clues for devising a new concept is, as we saw, valid only on the assumption that nature is amenable to uniformities. In this way, the assumption of uniformity is not only, or even primarily, a warrant for unification, but rather a basis for finding concepts to grasp the specificity and particularity of nature.

In focusing on reflective judgment rather than determinative judgment, Kant is dedicating his third *Critique* to the question of the relation between the particular and the universal, between object or intuition and concept, and thus between the lower and the higher faculties. As we will see in the following two subsections, this translates into a theory of the relation between the faculties that holds their mutual harmony to be a contingent occurrence rather than a universally guaranteed state. This is why we will now turn to the doctrine of the harmony of the faculties.

### 7.2.2 The Harmony of the Faculties and the Harmony of the Beautiful

Given that the *Critique of the Power of Judgment* is supposed to be a *Critique* of the feeling of pleasure and displeasure, which is the *Gemüt* insofar as the power of judgment is legislative in it, Kant needs to somehow connect judgment and pleasure. He does this in two seemingly separate manners. The first connection has to do with our intellectual pleasure in the attainment of our cognitive end of systematization:

The attainment of every aim is combined with the feeling of pleasure; and, if the condition of the former is an a priori representation, as in this case a principle for the reflecting power of judgment in general, then the feeling of pleasure is also determined through a ground that is a priori and valid for everyone; and indeed merely through the relation of the object to the faculty of cognition, without the concept of purposiveness in this case having the least regard to the faculty of desire, and thus being entirely distinct from any practical purposiveness of nature. (AA V: 187)

Here, the pleasure felt in attaining such a cognitive aim is supposed to be universal and a priori valid for everyone because the cognitive aim belongs to all those who have the

faculty of reason and are divided between sensibility and understanding. Unsurprisingly, Kant insists that what we experience here is the purposiveness of nature for our understanding, which is merely a relation of the object to the faculty of cognition, rather than a property of the object. This is important because Kant insists that we do not take the purposiveness in this sense to be an indicator for purposiveness insofar as it is considered in practical reason, i.e. as intentional. This corresponds with our analysis of chapter 4 and of 5.4.1, where the purposiveness of nature is in the first place not necessarily intentional, and in the second place appears intentional only to our cognitive capacities.

Kant insists that we feel such intellectual pleasure only if the task contains intentional effort and if it strikes us as non-obvious:

In fact, although in the concurrence of perceptions with laws in accordance with universal concepts of nature (the categories) we do not encounter the least effect on the feeling of pleasure in us nor can encounter it, because here the understanding proceeds unintentionally, in accordance with its nature, by contrast the discovered unifiability of two or more empirically heterogeneous laws of nature under a principle that comprehends them both is the ground of a very noticeable pleasure, often indeed of admiration, even of one which does not cease though one is already sufficiently familiar with its object. To be sure, we no longer detect any noticeable pleasure in the comprehensibility of nature and the unity of its division into genera and species, by means of which alone empirical concepts are possible through which we cognize it in its particular laws; but it must certainly have been there in its time, and only because the most common experience would not be possible without it has it gradually become mixed up with mere cognition and is no longer specially noticed. – It thus requires study to make us attentive to the purposiveness of nature for our understanding in our judging of it, where possible bringing heterogeneous laws of nature under higher though always still empirical ones, so that if we succeed in this accord of such laws for our faculty of cognition, which we regard as merely contingent, pleasure will be felt. Conversely, a representation of nature that foretold that even in the most minor investigation of the most common experience we would stumble on a heterogeneity in its laws that would make the unification of its particular laws under universal empirical ones impossible for our understanding would thoroughly displease us; because this would contradict the principle of the subjective-purposive specification of nature in its genera and our reflecting power of judgment with respect to the latter. (AA V: 187-188)

Kant states that there needs to be some sense of expectance that the process can go awry if we are to feel real intellectual pleasure in realizing our cognitive aims. That we do not usually feel such pleasure, is because we have either lost sight of the expectation of systematicity, or because we have lost sight of the contingency of the rational order

of nature. If we were to be reminded of that contingency, we would again marvel at the harmony of the empirical world.

Notoriously, though, Kant introduces a feeling of pleasure that has the same universal and a priori validity as that of intellectual pleasure, but does not stem from the satisfaction of a cognitive aim:

If pleasure is connected with the mere apprehension (*apprehensio*) of the form of an object of intuition without a relation of this to a concept for a determinate cognition, then the representation is thereby related not to the object, but solely to the subject, and the pleasure can express nothing but its suitability to the cognitive faculties that are in play in the reflecting power of judgment, insofar as they are in play, and thus merely a subjective formal purposiveness of the object. For that apprehension of forms in the imagination can never take place without the reflecting power of judgment, even if unintentionally, at least comparing them to its faculty for relating intuitions to concepts. Now if in this comparison the imagination (as the faculty of a priori intuitions) is unintentionally brought into accord with the understanding, as the faculty of concepts, through a given representation and a feeling of pleasure is thereby aroused, then the object must be regarded as purposive for the reflecting power of judgment. (AA V: 189-190)

Kant speaks here of a distinct feeling of pleasure that corresponds with the purposiveness of an object for the reflective power of judgment without there being a concept that is attained in reflection. There is thus a contrast with the abovementioned intellectual pleasure, which arises from our satisfaction in meeting a cognitive aim. In this case, however, no such cognitive aim is realized, and still we experience the object as purposive for the faculty of judgment. Kant states that this is possible because, in reflecting on such an object, an accord between the imagination and the understanding is unintentionally brought about. In this manner, Kant introduces his famous theory of the harmony of the faculties, which he characterizes in many ways throughout his work. In the first Introduction, the way in which the theory of the harmony of the faculties relates to that of aesthetic experience and that of reflective judgment is, as often, even more clearly put:

in the power of judgment understanding and imagination are considered in relation to each other, and this can, to be sure, first be considered objectively, as belonging to cognition (as happened in the transcendental schematism of the power of judgment); but one can also consider this relation of two faculties of cognition merely subjectively, insofar as one helps or hinders the other in the very same representation and thereby affects the state of mind, and [is] therefore a relation which is sensitive (which is not the case in the separate use of any other faculty of cognition). Now although this sensation is not a sensible representation of an object, still, because it is subjectively connected with the process of making the concepts of the understanding sensible by means of the power of judgment, it

can, as a sensible representation of the state of the subject who is affected by an act of that faculty, be reckoned to sensibility, and a judgment can be called aesthetic, i.e., sensible (as far as its subjective effect, not its determining ground is concerned), although judging (that is, objectively) is an action of the understanding (as the higher cognitive faculty in general) and not of sensibility. (AA XX: 223)

This passage is clearer because it immediately suggests how Kant believes to be able to solve the deadlock between rationalist and empiricist and between intellectualist and sensualist theories of aesthetic experience. In the first place, aesthetic experience can be a sensitive state rather than a cognitive one, because it is subjectively related to a state of the subject rather than objectively related to an object and the cognition thereof. But it has a clear connection with cognition, because it is connected with judgment rather than with sensations or inner senses, and is thus connected to an activity that is, strictly speaking cognitive. What is of interest to this cognitive activity is, however, not that or whether it attains its cognitive ends, but rather whether its subjective conditions are met, namely a mutual assistance, an accord, a harmony between the imagination and the understanding. In the official introduction, Kant also already suggests that this is enough to indicate why the feeling of pleasure felt on such an occasion is universal rather than merely particular:

someone who feels pleasure in mere reflection on the form of an object, without regard to a concept, rightly makes claim to the assent of everyone else, even though this judgment is empirical and is an individual judgment, since the ground for this pleasure is to be found in the universal though subjective condition of reflecting judgments, namely the purposive correspondence of an object (be it a product of nature or of art) with the relationship of the cognitive faculties among themselves (of the imagination and the understanding) that is required for every empirical cognition. The pleasure in the judgment of taste is therefore certainly dependent on an empirical representation, and cannot be associated a priori with any concept (one cannot determine a priori which object will or will not suit taste, one must try it out); but it is nevertheless the determining ground of this judgment only in virtue of the fact that one is aware that it rests merely on reflection and on the general although only subjective conditions of its correspondence for the cognition of objects in general, for which the form of the object is purposive. (AA V: 191)

Kant states here that this feeling is universal because it does not refer to some particularity about the constitution of the judging subject, but instead refers to the necessary subjective condition for all judging, namely the relation between the faculties. This means that we must now gain a better understanding of this relation between the faculties (in this subsection) and of the way in which that can be the ground for a feeling of pleasure (in the next subsection). The former task requires that

we not only explicitate the kind of relation between the faculties that is involved, but also how it can occur in the case of the judgment of a single object.

Kant's doctrine of the faculties departs from a central idea, namely that of a possible *disharmony* of the faculties – possible because of the conflict-admitting nature of their relation:

{*Taste* is art. The understanding and imagination, which have to unite in this, are like two friends, who cannot stand each other and yet cannot part from each other - for they live in perpetual strife and yet are mutually indispensable. [...] Imagination and understanding are two friends who cannot do without one another but cannot stand one another either, for one always harms the other. The more universal the understanding is in its rules, the more perfect it is, but if it wants to consider things *in concreto* then it absolutely cannot do without the imagination. (AA XXIV: 708-710)

In commenting on this passage, Henry Allison clearly recognizes the main thrust of Kant's doctrine:

the “friendship” between imagination and understanding is not without a certain tension, which results from the fact that they pull in opposite directions: the understanding toward universality and the imagination toward specificity. Accordingly, though the understanding requires the imagination to exhibit intuitively what is thought in its concept, and the imagination presumably needs the understanding to give it direction so that it can know what to exhibit, they nevertheless often work at cross purposes (and therefore “harm” one another). Although Kant does not spell it out, this presumably occurs either when the understanding in its endemic quest for universality produces a concept that is too general and indeterminate to be represented adequately *in concreto* by any particular instance, for example, the concept of a living thing, or when the particular imaginatively apprehended is too idiosyncratic or atypical to represent adequately what is thought in the concept, for example, the image of a three-legged dog. (Allison 2001: 48)

This reading is perfectly in line with the results of my previous two chapters. There, I claimed that reason allows for conflict because it has two distinct strivings, namely that towards generality and that towards particularity. In 6.3.1, I argued that, for Kant, understanding and sensibility, and more generally the higher and the lower faculties are distinct because they are the faculties for generality and particularity respectively. I have also already suggested that the faculties for generality and those for particularity are sanctioned by reason's drive for unity and for specificity respectively. All of this is reflected in the above-quoted passage, and in the general doctrine of the harmony of the faculties. As we will see later, it also resonates clearly with the organic picture of the Gemüt.



First however, we must turn to the experience that occasions our feeling of harmony amongst the faculties: aesthetic experience. According to Kant, in aesthetic experience, we are judging an object without the intention of deriving a concept from it. At some points, however, Kant insists that, in order for aesthetic appreciation to be distinct from sensual, moral or cognitive appreciation, no determinate concept may be used in aesthetic judgment (AA V: 229). This has puzzled interpreters because it is not clear whether Kant can accommodate such reasoning. After all, concepts need to be in play if judgment is to occur at all. However, it seems that Kant never goes so far as to deny that concepts can be at work in aesthetic experience: what he says is that we should have no determinate concept of the internal end of that object, of “the internal purpose that determines its possibility” (AA V: 230). This means that concepts may be employed, but not the concept that is meant to grasp the object in its individuality according to a principle. This corresponds with the analysis of reflective judgment in the previous subsection, where I proposed we read Kant as majorly concerned with devising an individuating concept, a concept with which to grasp the specificity of an object through conceptual means. The problem there was not that we cannot grasp the object with concepts, for we grasp the many features and properties it exhibits through concepts. But these concepts are not sufficient to see how these many properties unite into a individual in this case.

This reading also suggests itself from an analysis of Kant’s conception of what it is in the object that we perceive as beautiful, namely the “merely formal purposiveness”. In order to understand this, we need to contrast this formal purposiveness with material purpose. From the context of Kant’s discussion, the latter appears to correspond with the concept of perfection:

Objective purposiveness is either external, i.e., the **utility** of the object, or internal, i.e., its **perfection**. That the satisfaction in an object on account of which we call it beautiful could not rest on the representation of its utility is sufficiently obvious from the two preceding main sections, since in that case it would not be an immediate satisfaction in the object, which latter is the essential condition of the judgment about beauty. But an objective inner purposiveness, i.e., perfection, already comes closer to the predicate of beauty, and has therefore been held to be identical with beauty even by philosophers of repute, though with the proviso if **it is thought confusedly**. It is of the greatest importance in a critique of taste to decide whether beauty is really reducible to the concept of perfection. (AA V: 226-227)

In the light of my discussion from 7.1, it is clear whom Kant is targeting here: the Wolffian tradition and their conception of beauty as a sensible –and thus confused– cognition of the perfection of the object. In the previous chapter (6.3.2), we already saw that Kant took issue with the metaphysical concept of perfection as internal unity of diversity. This also resonates with Hutcheson, who similarly believed that the sense of

beauty is the inner sensation corresponding with the perfection of the object. Given this definition of perfection as unity of the diverse, we can see that Kant sought to repudiate the idea of an aesthetic experience as a recognition, however confusedly or indirectly, of how an object, in all its specificity and unity, exhibits a clear conceptual unity:

To judge objective purposiveness we always require the concept of an end, and [if that purposiveness is not to be an external one (utility), but an internal one], we require the concept of an internal end, which contains the ground of the internal possibility of the object. Now as an end in general is that the **concept** of which can be regarded as the ground of the possibility of the object itself, thus in order to represent an objective purposiveness in a thing the concept of **what sort of thing it is supposed to be** must come first; and the agreement of the manifold in the thing with this concept (which supplies the rule for the combination of the manifold in it) is the **qualitative perfection** of a thing. (AA V: 227)

In understanding an object as objectively purposive, we understand through the concept of what it is supposed to be, i.e. through the rule that governs the unity of the manifold. In such a judgment, both the kind and the individuality of the object are understood as unified in a specific manner, as constituting, I submit, something like an essence, rather than a mere aggregate of properties. We already encountered one form of thinking on objective purposiveness, namely in the case of teleology: Kant believed that natural purposes are those objects which need to be understood through their internal purpose, through the unified kind or individuality that makes them what they are and properly grasp their peculiar unification in diversity. But throughout the Critique, as we already saw in 4.2.1, Kant sought to strictly separate objective and subjective purposiveness.

This raises the question regarding the nature of subjective or formal purposiveness, a question that is all the more vexing because Kant seems to waver between two opinions. The first opinion is that formal or subjective purposiveness has little or nothing to do with the structure of the object itself, and only with its relation to the subject:

What is formal in the representation of a thing, i.e., the agreement of the manifold with a unity (leaving undetermined what it is supposed to be) does not by itself allow any cognition of objective purposiveness at all, because since abstraction is made from this unity, **as an end** (what the thing is supposed to be), nothing remains but the subjective purposiveness of representations in the mind of the beholder, which indicates a certain purposiveness of the representational state of the subject, and in this an ease in apprehending a given form in the imagination, but not the perfection of any object, which is here not conceived through any concept of an end. (AA V: 227)

In the Dohna Wundlacken-Logic, this idea is condensed into the contrast between *logical* and *aesthetic* purposiveness:

1. logical perfection in the agreement of the faculty of cognition with the object,
2. aesthetic perfection[;] it consists in the subject's faculty of cognition. (AA XXIV: 705)

These passages suggest that beauty is not in the object, but in the mental state of the beholder, or rather judge. However, on other occasions Kant does seem intent on identifying the formal purposiveness of the beautiful object as a property of the representation of that object, and not just as a representation of a state of the subject. This raises the following question: what would it mean for an object to be judged without a concept and yet appear as having certain, albeit formal characteristics. If we want to overcome this difficulty, we need to see what properties a representation of an object would have to exhibit if it were to prompt the state in which its aesthetic appreciation consists.

A first important clue can be gathered from the following passage from the “General remark on the first section of the *Analytic*”:

If one draws the conclusion from the above analyses, it turns out that everything flows from the concept of taste as a faculty for judging an object in relation to the **free lawfulness** of the imagination. But if in the judgment of taste the imagination must be considered in its freedom, then it is in the first instance taken not as reproductive, as subjected to the laws of association, but as productive and self-active (as the authoress of voluntary forms of possible intuitions); and although in the apprehension of a given object of the senses it is of course bound to a determinate form of this object and to this extent has no free play (as in invention), nevertheless it is still quite conceivable that the object can provide it with a form that contains precisely such a composition of the manifold as the imagination would design in harmony with the **lawfulness of the understanding** in general if it were left free by itself. Yet for the **imagination** to be free and yet **lawful by itself**, i.e., that it carry autonomy with it, is a contradiction. The understanding alone gives the law. But when the imagination is compelled to proceed in accordance with a determinate law, then how its product should be, as far as its form is concerned, is determined through concepts; but then, as was said above, the satisfaction is not that in the beautiful, but in the good (of perfection, in any case merely the formal kind), and the judgment is not a judgment by means of taste. Thus only a lawfulness without law and a subjective correspondence of the imagination to the understanding without an objective one – where the representation is related to a determinate concept of an object – are consistent with the free lawfulness of the understanding (which is also called purposiveness without an end) and with the peculiarity of a judgment of taste. (AA V: 240-241)

Like so many passages in Kant, this one is enlightening because of its negative content: it tells us what aesthetic experience does not consist in. Kant insists that we experience beauty in an object when, in the act of judging it, imagination exhibits a free

lawfulness. Imagination is thus to exhibit some lawfulness, point to some unity, although it may not be constrained to exhibit this unity. This gives us an important clue as to the nature of the object that we experience as beautiful: it needs to appear to us, in judging, as a unity without there being a concept that determines this unity. As we saw above, perfection is a judgment that an object's diversity is conceptually integrated in an intimate way, i.e. that there seems to be a ruling principle governing that object. In the case of an organism, such an object seemed to instantiate a concept as a rule or norm for its functional integration. An aesthetically perfect object, I submit, shows formal integration rather than material integration, i.e. shows a unity of diversity that is not recognized through or reducible to a governing principle or rule. I will abstain here from an elaboration of Kant's theory of beauty and art, and explaining my own conviction that art is to be considered as non-conceptual integration, but will simply show how this reading allows us to understand and qualify some central tenets of Kant's philosophy, namely his formalism and his anti-classicism.

Kant is known as an unflinching formalist in aesthetics, since he defends the idea that only the form of the object is concerned in aesthetic experience. We have already seen what this means in one sense, namely that its logical perfection should not be concerned. But Kant's theory of "disinterestedness" also entails that, besides objective internal purposiveness, external purposiveness like that of the sensually agreeable and the morally desirable are not at stake. The reason why he stresses this is because he wants to discuss aesthetic experience as autonomous from other kinds of experience. The objective internal purposiveness is the proper concern of the faculty of cognition, and the external purposiveness is the concern for the faculty of desire. Kant's question regarding the possibility of aesthetic judgment is thus also a question as to the possibility of a level of appreciation autonomous from the cognitive and the moral. He thereby betrays his dissatisfaction with those accounts which sought to assimilate the aesthetic to one of these other functions –not because in doing so they are straightforwardly mistaken, but rather because they ultimately deny the possibility of aesthetic experience.

Many readers have suggested that this commitment to the autonomy and purity of the aesthetic led Kant astray in his theory, because it led him to unduly restrict the scope of the aesthetic. This judgment, however, may be unfair to Kant. First of all, it needn't be so that Kant does not allow for such "material" elements to enter into consideration – he merely requires that they are not the basis for the aesthetic judgment. In its turn, this needn't imply, as some seem to suggest, that the material elements should be left out of consideration in aesthetic judgment completely. On such a reading, however, it is not clear of what the form is the form. By the merely formal, Kant needn't mean shape in the straightforward spatial sense. After all, Kant denies that it is spatial configuration with which we are primarily concerned in aesthetic judgment, spatial configuration being primarily understood through mathematics.

I suggest that what Kant means when he speaks of the merely formal, is the manner of interrelation and integration of the material elements. Their mutual relation might be understood on the basis of their cognitive, agreeable or moral features, but the pleasure in the aesthetic does not consist in these features: it consists in how they relate, contrast, meet and merge into a complex whole, and that these relations are not themselves grasped as agreeable, morally desirable or cognitively adequate. Rachel Zuckert has offered a similar interpretation:

Thus, I have argued, Kant's aesthetic formalism comprises the claim that in appreciating an object as beautiful, we represent it as an individual, as a unity of diverse properties, reciprocally contrasting and complementary, or as formally purposive without a purpose. This representation is merely subjective (ly purposive) because it is not unified in accord with a concept, indeed cannot be characterized by discursive conceptualization; it cannot be or ground an objective judgment. (Zuckert 2007: 330)

I believe Kant insists that we do not think of beautiful objects as conceptually unified in another way as well. He insists, time and time again, that in aesthetics the imagination is to be free, and cannot simply be following an external rule. In the text, he is most concerned with stressing its independence from the lawfulness and the rules of the understanding. In his aesthetic theory, this first translates into his distaste of the classicist observance of rules in composition, disposition and execution of artworks:

In a thing that is possible only through an intention, in a building, even in an animal, the regularity that consists in symmetry must express the unity of the intuition, which accompanies the concept of the end and belongs to the cognition. But where only a free play of the powers of representation (although under the condition that the understanding does not thereby suffer any offense) is to be maintained, in pleasure gardens, in the decoration of rooms, in all sorts of tasteful utensils and the like, regularity that comes across as constraint is to be avoided as far as possible; hence the English taste in gardens or the baroque taste in furniture pushes the freedom of the imagination almost to the point of the grotesque, and makes this abstraction from all constraint by rules the very case in which the taste can demonstrate its greatest perfection in projects of the imagination.

All stiff regularity (whatever approaches mathematical regularity) is of itself contrary to taste: the consideration of it affords no lasting entertainment, but rather, insofar as it does not expressly have cognition or a determinate practical end as its aim, it induces boredom. (AA V: 242-243)

Kant thus insists that art governed by concepts, by rules, by considerations of regularity, fails to please precisely because of its clear unity. In such works, the unity is seen as too laboured, as smelling of the lamp, and moreover it is seen as lacking in

novelty and originality. This judgment on the formulaic equally lies behind Kant's theory of genius:

genius 1) is a **talent** for producing that for which no determinate rule can be given, not a predisposition of skill for that which can be learned in accordance with some rule, consequently that **originality** must be its primary characteristic. 2) That since there can also be original nonsense, its products must at the same time be models, i.e., **exemplary**, hence, while not themselves the result of imitation, they must yet serve others in that way, i.e., as a standard or a rule for judging.<sup>a</sup> 3) That it cannot itself describe or indicate scientifically how it brings its product into being, but rather that it gives the rule as **nature**, and hence the author of a product that he owes to his genius does not know himself how the ideas for it come to him, and also does not have it in his power to think up such things at will or according to plan, and to communicate to others precepts that would put them in a position to produce similar products. (For that is also presumably how the word "genius" is derived from *genius*, in the sense of the particular spirit given to a person at birth, which protects and guides him, and from whose inspiration those original ideas stem.) 4) That by means of genius nature does not prescribe the rule to science but to art, and even to the latter only insofar as it is to be beautiful art. (AA V: 307 – 308)

Genius is thus that which allows one to create wonderfully integrated wholes without explicitly considering the integrating relations, without at each time seeking them out on the basis of unified concept. This does not mean that no awareness can be present in the activity of the genius, but only that, in the end, the harmony and integration escapes even him. Furthermore, the integration of a work of genius can later be conceptually articulated, can be studied, i.e. can be study of what we now call criticism. But the goal of criticism is not to understand what makes something beautiful: it is to understand what made *this* beautiful, and what new options, what new free creations can be inspired by it. The connection with this theory and that of the free harmony of the faculties wherein the imagination is freely active is explicit in the third *Critique*:

The mental powers, then, whose union (in a certain relation) constitutes **genius**, are imagination and understanding. Only in the use of the imagination for cognition, the imagination is under the constraint of the understanding and is subject to the limitation of being adequate to its concept; in an aesthetic respect, however, the imagination is free to provide, beyond that concord with the concept, unsought extensive undeveloped material for the understanding, of which the latter took no regard in its concept, but which it applies, not so much objectively, for cognition, as subjectively, for the animation of the cognitive powers, and thus also indirectly to cognitions[.] (AA V: 316-317)

But this brings us back to the harmony of the faculties. In the passage where Kant speaks of the freedom of the imagination, he is not just saying that imagination should

be free from the understanding – he also insists that it must be independent from the senses and from the law of association. This is more important than may seem at first, because Kant does believe that we cannot a priori determine whether a given object will be perfect: we need to experience it, judge it, in order to appreciate it. This suggests that in experiencing an object, it is the receptive element that plays the important role. But this is not what Kant says – he insists that what is at stake is the power of the imagination, not the givenness of sensibility.

An initial answer might be suggested by remembering Edmund Burke's Lockean distinction between imagination and reasoning: imagination seeks and suggests similarities through association. Reasoning is primarily concerned with differentiating. This picture has initial plausibility as a framework for understanding Kant, because it describes the imagination and the understanding as engaged in opposite activities. The harmony of the faculties would then consist in the correspondence between the associations suggested by the imagination and the conceptual distinctions introduced by the understanding. This interpretation cannot stand, however, since Kant insists that the power of the imagination that it freely employs in its play is not that of association: for Kant, association is not free at all, since it is based on the reproductive aspect of imagination. Moreover, this picture inverts Kant's division of labour between the two faculties: for him, it is understanding that insists on similarities, whereas imagination serves variety and specificity. Thus, the associative use of the imagination is in service of the understanding, and not a free action of the imagination itself.

This might all raise questions as to what talk of the harmony of the faculties is meant to express. In the face of Kant's tendency to express his ideas in terms of faculties rather than in terms of the objects to which they are related, has frustrated many readers. Their frustration, however, might simply be the result from their unwillingness to take the talk of the harmony of the faculties as more informative, namely by taking the various concepts to be not vague suggestions, but real conceptual articulations.

In the passage on which I have been elaborating, Kant suggests that, in the experience of the beautiful, the imagination, taken as a distinct, active faculty of the mind, with its own proper contributions to knowledge, reveals itself as in harmony rather than in tension with another faculty, namely the understanding. It stands in this harmony not because there is a transcendent or external principle which causes it to work in this way. On the contrary, its own "productive and self-active" is already, of itself, functioning so as to maximally service the understanding. The aesthetic experience, thus seems to indeed be of a state of the subject: it is, as I will argue in the next-subsection, the experience of the functional unity in diversity of the system of faculties itself.

### 7.2.3 The Pleasure in Harmony

In the previous subsection, I mentioned that Kant usually speaks of the experience of the beautiful as concerning a state of the subject rather than a property of an object. I have tried to indicate how this is compatible with the idea that a beautiful object nonetheless exhibits some features which entitle it to be called beautiful, and that beauty is not in any straightforward sense merely in the eye of the beholder. Nevertheless, the beauty of the object plays a peculiarly peripheral role in Kant's aesthetics. Many readers have been frustrated by this, since instead of accounts of what he means by the "form" of the object insofar as it is aesthetically relevant, Kant keeps insisting on a peculiar doctrine of the harmony of the faculties. In this study, however, it is the theory of the faculties that is of more interest than the theory of art, and hence we should look whether there is some convergence between the various metaphorical expressions of the doctrine of harmony. I will of course argue that there is such a convergence, namely in that they all refer to the organic picture of the mind first introduced in 3.4, and that they indicate how this organic unity is *normative* in the sense of chapter 4.

One of the most crucial and most puzzling characterizations of the harmony of the faculties is from §35 of the *Critique of the Power of Judgment*:

The subjective condition of all judgments is the faculty for judging itself, or the power of judgment. This, employed with regard to a representation by means of which an object is given, requires the agreement of two powers of representation: namely, the imagination (for the intuition and the composition of the manifold of intuition), and the understanding (for the concept as representation of the unity of this composition). Now since no concept of the object is here the ground of the judgment, it can consist only in the subsumption of the imagination itself (in the case of a representation by means of which an object is given) under the condition that the understanding in general advance from intuitions to concepts. I.e., since the freedom of the imagination consists precisely in the fact that it schematizes without a concept, the judgment of taste must rest on a mere sensation of the reciprocally animating imagination in its **freedom** and the understanding with its **lawfulness**, thus on a feeling that allows the object to be judged in accordance with the purposiveness of the representation (by means of which an object is given) for the promotion of the faculty of cognition in its free play; and taste, as a subjective power of judgment, contains a principle of subsumption, not of intuitions under **concepts**, but of the **faculty** of intuitions or presentations (i.e., of the imagination) under the **faculty** of concepts (i.e., the understanding), insofar as the former **in its freedom** is in harmony with the latter **in its lawfulness**. (AA V: 287)



What Kant expresses here is that the aesthetic feeling is more concerned with the act of judging than with the judged object, and that the subjective state to which the experience of beauty refers is the state of the faculties such that the act of judgment is facilitated. Moreover, the passage is explicit in noting that the judged object is only relevant as occasioning such an accord or harmony. The harmony of the faculties is therefore characterized as an act of judgment, i.e. an act of subsumption, but not of a particular under a concept, but rather of the capacity for presenting particulars under the capacity to think through concepts. Unfortunately, this talk of subsumption is of course even more metaphorical, since two faculties cannot be subsumed under each other in the strict sense: imagination is not itself an intuition, and understanding is not itself a concept. Therefore, it is best to read Kant as saying that the faculties are in harmony when imagination is felt to generally present particulars of its own accord in such a way that they lend themselves to being conceptually grasped without belaboured and undue abstractions. What we experience in the harmony is, then, not the success of our judgment in this case, but rather a state in which judgment would generally be successful.

This reveals why Kant introduces a second characterizing of pleasure when he starts dealing with aesthetic judgment. In the previous subsection, we saw that the accomplishment of an aim prompts pleasure. Now, he also introduces the following:

The consciousness of the causality of a representation with respect to the state of the subject, **for maintaining** it in that state, can here designate in general what is called pleasure; in contrast to which displeasure is that representation that contains the ground for determining the state of the representations to their own opposite (hindering or getting rid of them). (AA V: 220)

This general contrast between pleasure and pain (Lust und Unlust) reveals that pleasure prompts us to retain a given state, whereas pain prompts us to leave it. In the case of the agreeable, the sensually pleasurable, we are prompted to keep enjoying that state, for instance by seeking renewed contact with the pleasurable object. In aesthetic judgment, we are prompted to maintain the specific alignment of faculties realized in the act of judging a particular object. A definition of pleasure from the *Critique of Practical Reason* allows us to better understand the connection between these two kind of pleasure, and why we are prompted, in both cases, to maintain our state:

The faculty of desire is a being's *faculty to be by means of its representations the cause of the reality of the objects of these representations*. Pleasure is the *representation of the agreement of an object or of an action with the subjective conditions of life*, i.e., with the faculty of the *causality of a representation with respect to the reality of its object* (or with respect to the determination of the powers of the subject to action in order to produce the object). (AA V: 9)

Here, Kant defines pleasure as “the representation of the agreement of an object or of an action with the subjective conditions of life”. We cannot straightforwardly transpose this to the case of the third *Critique*, however, since Kant is here mainly concerned with the pleasurable as the agreeable, and as considered by the faculty of desire. But it is possible to see what changed between the *Critique of Practical Reason* and the *Critique of the Power of Judgment*. In the second *Critique*, the assumption is still that the object is connected so intensely with the state of the subject that we are interested in its reality and its realization. In the third *Critique*, however, Kant introduces a kind of purposiveness of the object for the subject that is less strictly connected, such that it is the subjective conditions that are the primary concern. In aesthetic judgment, we do not mean to preserve our state by preserving the object, we mean to preserve our state by preserving our subjective conditions directly.

In offering this reading, I may be taken to overlook another major difference between the two cases, namely that in the second *Critique* passage, Kant is speaking of the subjective conditions of life, whereas in the third *Critique*, he is speaking of the subjective conditions of *judgment*, i.e. *cognitive activity*. However, this is not an oversight, since I precisely mean to suggest that there is a greater parallel here as well. At the close of the previous chapter, I argued that Kant saw the system of faculties as somehow analogous to an organism in its complex functional unity. This means that the system of faculties has an analogue of *life*, its primary function being *cognitive activity*. If we accept this, Kant’s doctrine of the harmony of the faculties becomes more readily understandable: the agreeable is pleasurable because it promotes or agrees with the subjective conditions of our empirical, biological life, whereas the beautiful is pleasurable because it promotes or agrees with the subjective conditions of our cognitive activity, the proper arrangement and alignment of the faculties cooperating in *judgment*.

This makes sense of much of the peculiarities of Kant’s account of aesthetics, namely of how it can postulate a principle which is subjective and universally valid. According to Kant, this is only possible if the aesthetic judgment is somehow *pure*. However, it is difficult to understand how a state involving pleasure can be pure. We can understand this if we make a distinction between two kinds of pleasure, namely the pleasure that we find in the agreeable, and that flows from our biological and psychological needs, and the pure pleasures and feelings that stem from, and concern, the cognitive and moral needs of the *Gemüt*. However, although Kant makes such a distinction, he does not locate aesthetic pleasure in either of these categories:

That pleasure which is necessarily connected with desire (for an object whose representation affects feeling in this way) can be called *practical pleasure*, whether it is the cause or the effect of the desire. On the other hand, that pleasure which is not necessarily connected with desire for an object, and so is not at bottom a pleasure in the existence of the object of a representation but is attached only to

the representation by itself, can be called merely contemplative pleasure or *inactive delight*. We call feeling of the latter kind of pleasure *taste*. Practical philosophy, accordingly, speaks of contemplative pleasure only *in passing*, not as if the concept *belonged within* it. As for practical pleasure, that determination of the faculty of desire which is caused and therefore necessarily *preceded* by such pleasure is called *desire* in the narrow sense; habitual desire is called *inclination*; and a connection of pleasure with the faculty of desire that the understanding judges to hold as a general rule (though only for the subject) is called an *interest*. So if a pleasure necessarily precedes a desire, the practical pleasure must be called an interest of inclination. But if a pleasure can only follow upon an antecedent determination of the faculty of desire it is an intellectual pleasure, and the interest in the object must be called an interest of reason; for if the interest were based on the senses and not on pure rational principles alone, sensation would then have to have pleasure connected with it and in this way be able to determine the faculty of desire. Although where a merely pure interest of reason must be assumed no interest of inclination can be substituted for it, yet in order to conform to ordinary speech we can speak of an inclination for what can be an object only of an intellectual pleasure as a habitual desire from a pure interest of reason; but an inclination of this sort would not be the cause but rather the effect of this pure interest of reason, and we could call it a *sense-free inclination (propensio intellectualis)*. (AA VI: 212-213)

In this passage from the preface of the *Metaphysics of Morals*, a work published many years after the *Critique of the Power of Judgment*, Kant distinguishes sensual and intellectual pleasures. Presumably he needs the latter category to explain why the moral law can be action-motivational, through the pure feeling of awe, without making moral action impure by the admixture of a feeling. Note, however, that Kant does not locate aesthetic pleasure in either of these categories: he suggests that it properly belongs to neither, because it is not connected with practical, in other words, with the capacity to act in accordance with ends. This may make it troubling to understand exactly how the feeling of the beautiful can be connected with a striving to maintain a state, since no intentional action is involved. The organic analogy, however, allows us to avoid this conclusion. In chapter 4, we saw that Kant did not believe that organisms strive according to ends – their striving should thus properly be understood as non-intentional. I submit therefore that the striving involved in the harmony of the faculties is analogous to that of the unintentional striving that we see at work in plants and other non-animal life. This reading also makes sense of the following passage, because it suggests that aesthetic feeling is like moral and intellectual pleasure in being pure and stemming from an interest of reason rather than from a biological or psychological need, but unlike it in that it is practical-intentional:

To establish *a priori* the connection of the feeling of a pleasure or displeasure as an effect with some representation (sensation or concept) as its cause is absolutely

impossible, for that would be a causal relation, which (among objects of experience) can only ever be cognized *a posteriori* and by means of experience itself. To be sure, in the critique of practical reason we actually derived the feeling of respect (as a special and peculiar modification of this feeling, which will not coincide exactly either with the pleasure or with the displeasure that we obtain from empirical objects) from universal moral concepts *a priori*. But there we could also step beyond the bounds of experience and appeal to a causality that rests on a supersensible property of the subject, namely that of freedom. But even there we did not actually derive this **feeling** from the idea of the moral as a cause, rather it was merely the determination of the will that was derived from the latter. The state of mind of a will determined by something, however, is in itself already a feeling of pleasure and is identical with it, thus it does not follow from it as an effect: the latter would only have to be assumed if the concept of the moral as a good preceded the determination of the will by the law, for in that case it would be pointless for the pleasure that would be connected with the concept to be derived from it as a mere cognition.

Now it is similar with the pleasure in the aesthetic judgment, except that here it is merely contemplative and does not produce an interest in the object, while in the moral judgment it is practical. The consciousness of the merely formal purposiveness in the play of the cognitive powers of the subject in the case of a representation through which an object is given is the pleasure itself, because it contains a determining ground of the activity of the subject with regard to the animation of its cognitive powers, thus an internal causality (which is purposive) with regard to cognition in general, but without being restricted to a particular cognition, hence it contains a mere form of the subjective purposiveness of a representation in an aesthetic judgment. (AA V: 221-222)

This passage is also interesting for two other reasons. The first is that Kant clarifies his causal language in his aesthetic theory. Paul Guyer (1997: 93-97) has argued that we need to take this causal language literally, and therefore assume that the aesthetic feeling is a matter of psychological mechanisms proper to the minds of men. In such a case, the dynamics of the faculties described here is not between the faculties as considered by epistemology (as in the first *Critique*), but between the faculties considered as psychological and empirical. This conclusion does not need to follow however, since in the passage Kant quoted above he denies that the sense in which the moral law determines the finite agent to act morally is properly *causal* in the sense of the empirical category. Instead, this determination is to be regarded as some internal determination that is analogous to what we would improperly call a causality. I suggest that, when Kant says in that passage that the aesthetic feeling is similar to the moral, it is also in being determined suprasensibly rather than causally. This does not quell the problems of transcendental idealism, but it also does not raise problems that were not already contained in the idea of moral determination.

The second interesting aspect about the passage is its characterization of the harmony of the faculties as an “animation” (Belebung) of the faculties. The same characterization is also to be found elsewhere, for instance in the following passages:

The animation (Belebung) of both faculties (the imagination and the understanding) to an activity that is indeterminate but yet, through the stimulus of the given representation, in unison... (AA V: 219)

the facilitated play of both powers of the mind (imagination and understanding), enlivened (belebt) through mutual agreement. (AA V: 219)

For readers like Guyer (1997: 97), this language of animation and livening is too figurative. I agree that, given the fact that Kant is speaking of transcendental faculties rather than of psychological or biological faculties, the language is to a certain extent figurative. I do not agree, however, that it is therefore devoid of conceptual portent. If we read Kant as conceiving of the *Gemüt* as an organic unity of distinct and possibly uncooperative faculties, then we can gain a better understanding of the way in which the faculties “animate” each other. In 6.2.2, we saw that Crusius explicitly spoke of the vivification or animation of faculties by other faculties. This meant that certain faculties were further developed or activated by a corresponding activity of another fundamental faculty. A similar picture seems to lie behind Kant’s use of that language here. This would mean that Kant saw the harmony of the faculties as one which is desirable for the system of the faculties because in it the different capacities are strengthened and led to express themselves. But, as I suggested already in 5.4, in the mature Kant, the dynamical metaphor of “life” as in “living force” yields to that of “life” as in “living system”. This results in the following reading: in the harmony of the faculties, the different faculties of the mind, which are like the different activities of different organs, influence each other, strengthen each other and specify in such a way that the unity of the *Gemüt* is thereby reinforced. It is because in such a state the *Gemüt*, like an organism, appears as apt to its primary function, namely cognition, without there having to be a leading or transcendent principle to make it apt in this way, that there is a striving to maintain this functional state.

If we take all this together, we can answer one of the two questions with which I ended chapter 6.4. There, I inquired whether the system of faculties is like an organism not only because it is a unity of diversity in the relevant sense, but also because it is equally open to contingency. In offering his theory of the harmony of the faculties, Kant seems to be saying that the maximal unity of the system of faculties is not always guaranteed, and is in fact contingent upon certain experiences, certain concrete and contingent cases of judgment. Whether our system of faculties is internally aligned in such a way that it allows us to judge almost effortlessly, is something which only

experience can teach us. When, in our experiencing or judging, this maximal unity is felt or realized, the system of faculties is determined to maintain this state.

What remains now is to answer the other question, namely whether besides being truly normative and open to contingency, the system of faculties is also plastic and epigenetic. This question will be dealt with in the following section.

### 7.3 The Sublime Vocation of Man

Kant does not limit his aesthetics to a discussion of the beautiful, for he also speaks of the sublime. In this, he is most likely inspired by Burke. According to Burke, the beautiful is correlated with pleasure, whereas the sublime is correlated with some displeasure, which nonetheless prompts some delight. Kant offers the same general picture:

the latter pleasure is very different in kind from the former, in that the former (the beautiful) directly brings with it a feeling of the promotion of life, and hence is compatible with charms and an imagination at play, while the latter (the feeling of the sublime) is a pleasure that arises only indirectly, being generated, namely, by the feeling of a momentary inhibition of the vital powers and the immediately following and all the more powerful outpouring of them; hence as an emotion it seems to be not play but something serious in the activity of the imagination. Hence it is also incompatible with charms, and, since the mind is not merely attracted by the object, but is also always reciprocally repelled by it, the satisfaction in the sublime does not so much contain positive pleasure as it does admiration or respect, i.e., it deserves to be called negative pleasure. (AA V: 244-245)

Kant's doctrine of the sublime is interesting for several reasons. First of all, it is of course tantamount to the claim that not everything that is not beautiful is thereby either ugly or aesthetically uninteresting. Even a negative feeling, Kant seems to say, can be the basis of a pleasure, in which case "it deserves to be called negative pleasure". The second reason is that Kant again connects the feeling of the sublime with the organic and vital metaphors which were so dominant in his account of the beautiful. Here too, Burke lies in the background. On Burke's physiological aesthetics, the sublime originates in experiences of fear, prompted by that which threatens our life, but which turns out to be harmless in spite of its terribleness. In light of the analysis of 6.4 and

7.2.3, we can read Kant's theory of the sublime as a transcendentalization of Burke's merely physiological account.<sup>4</sup> On such a reading, the sublime is the experience of an object which is in the first order harmful to the harmony of the faculties, but in the second order introduces another sense of adaptedness. In this section, I will attempt to articulate such a reading, rather than give a complete account of Kant's theory of the sublime

Kant continues to contrast the beautiful and the sublime as follows:

The most important and intrinsic difference between the sublime and the beautiful, however, is this: that if, as is appropriate, we here consider first only the sublime in objects of nature (that in art is, after all, always restricted to the conditions of agreement with nature), natural beauty (the self-sufficient kind) carries with it a purposiveness in its form, through which the object seems as it were to be predetermined for our power of judgment, and thus constitutes an object of satisfaction in itself, whereas that which, without any rationalizing, merely in apprehension, excites in us the feeling of the sublime, may to be sure appear in its form to be contrapurposive for our power of judgment, unsuitable for our faculty of presentation, and as it were doing violence to our imagination, but is nevertheless judged all the more sublime for that. (AA V: 245)

The first-order displeasure of the sublime is thus due to the fact that the "object" is unpurposive for our faculty of judgment because it is "doing violence to our imagination". The discussion of the beautiful allows us to understand why Kant singles out imagination as the primary victim. In the case of the beautiful, the imagination succeeds in schematizing and presenting the manifold in such a way that it suggests an integrated unity without having to employ the concept or principle of such a unity. We can thus expect that in the case of the sublime, the manifold is presented in such a way that the imagination seems to be in principle incapable of presenting it as unified:

in that which we are accustomed to call sublime in nature there is so little that leads to particular objective principles and forms of nature corresponding to these that it is mostly rather in its chaos or in its wildest and most unruly disorder and devastation, if only it allows a glimpse of magnitude and might, that it excites the ideas of the sublime. (AA V: 246)

But if this is true, then why do we nevertheless take some delight in the sublime? Why do we choose to revel so in the limitations, the inaptitude of our own cognitive organization? In order to answer this, Kant first insists that we are never presented with the sublime. This should be unsurprising, since he believes that we can only ever be

---

<sup>4</sup> Kant himself suggests this on AA V: 277.

presented with something if there is some conceptual grasp on it. If the sublime is that which is in principle resistant to such a conceptual grasp, then it is impossible for us to recognize something as sublime. Kant admits that much and therefore advances the following bold claim: “That is sublime which even to be able to think of demonstrates a faculty of the mind that surpasses every measure of the senses”. (AA V: 250)

Kant’s theory of the sublime is thus that, in experiencing, in aesthetic judgment, the inaptness of our imagination, we thereby nonetheless discover our own capacity to think that which lies beyond the limits of our capacities. This capacity is, of course, reason. For that reason, Kant describes the sublime as pertaining not to the relation between imagination and understanding, but rather between imagination and reason. What we discover, however, is not the infinite capacity of reason, but rather reason’s demand that the other faculties seek to overcome their limitations:

The feeling of the inadequacy of our capacity for the attainment of an idea **that is a law for us** is **respect**. Now the idea of the comprehension of every appearance that may be given to us into the intuition of a whole is one enjoined on us by a law of reason, which recognizes no other determinate measure, valid for everyone and inalterable, than the absolute whole. But our imagination, even in its greatest effort with regard to the comprehension of a given object in a whole of intuition (hence for the presentation of the idea of reason) that is demanded of it, demonstrates its limits and inadequacy, but at the same time its vocation for adequately realizing that idea as a law. Thus the feeling of the sublime in nature is respect for our own vocation, which we show to an object in nature through a certain subreption (substitution of a respect for the object instead of for the idea of humanity in our subject), which as it were makes intuitable the superiority of the rational vocation of our cognitive faculty over the greatest faculty of sensibility. (AA V: 257)

As mentioned above, Kant denies that any object can be truly presented as sublime. Now he explains how we nevertheless come to judge it as sublime, namely through a subreption. Strictly speaking, he says, it is not the object that is experienced and judged as sublime: it is the idea of humanity in ourselves. In his treatment of the sublime, Kant Kant systematically refers to the idea of humanity as a *vocation*. As a result, it is the vocation of man more than anything else which is sublime, i.e. the most awesome and the most terrible<sup>5</sup>:

---

<sup>5</sup> This resonates with Heidegger’s famous characterization of man as to deinotaton in *Introduction to Metaphysics* (2000: 160), because he is at once in nature and limited by nature, and capable of resisting and overcoming nature.



No idea so elevates the human mind and animates it even to inspiration as that of a pure moral disposition, revering duty above all else, struggling with the countless ills of life and even with its most seductive allurements and yet overcoming them (as we may rightly assume that one is capable of doing). That the human being is aware that he can do this because he ought to discloses within him a depth of divine predispositions and lets him feel, as it were, a holy awe at the greatness and sublimity of his true vocation. (AA VIII: 287)

We have already encountered the idea of man as vocation earlier on in this dissertation, namely in the third chapter, where I discussed Kant's theory of history. By integrating my conclusions from section 3.2 into the present context, we can arrive at the conclusion that the experience of the sublime opens man up to his historicity and his perfectibility.

In 3.2, I argued that Kant introduced his conception of humanity as a vocation as an attack on Rousseau's idea of faculty fixism. I characterized faculty fixism as the doctrine that man has well-defined epistemic capacities, which correspond to his essence, and that he can make no coherent attempt to change himself. Kant wished to oppose this idea and defend the cultural perfectibility and the historical malleability of man. The faculty that allows us to be so historically malleable and perfectible, Kant believed, is reason: the capacity to think the absolute, to think the extension beyond the current limitation, and to act morally. In the *Critique of the Power of Judgment*, Kant now argues that what allows us to experience man as a vocation rather than as a brute given is the experience of the sublime.

The idea that the "vocation" of man is to be read as the alterability, the perfectibility, the plasticity of human nature is further supported by the fact that, on two occasions (AA V: 247; AA V: 258) Kant states that the beautiful brings about a state of calm, whereas the sublime brings about movement. That the beautiful leads to calm, is to be expected, because it prompts a state which the *Gemüt* wants to maintain. The sublime, however, brings about a specific kind of movement that Kant calls a vibration, because we are at once repelled and attracted to the object. I suggest we read this as follows: in the second order, the sublime puts us in a state of displeasure, i.e. of wanting to leave the current state. But because we are at the same time attracted to the object, we cannot leave the state by avoiding the object. Hence, the only way to leave the state is to recalibrate, realign and expand our faculties.

That Kant is primarily concerned, in his discussion of the sublime, with the perfectibility of man, and with not accepting brute limits on our capacities, appears also from a tirade included in the "general remark on the exposition of aesthetic reflective judgments":

a religious sermon that preaches a groveling, base currying of favor and self-ingratiation, which abandons all confidence in our own capacity for resistance against evil, instead of the energetic determination to seek out the powers that

still remain in us, despite all our frailty, for overcoming inclinations; the false humility that finds the only way to be pleasing to the supreme being in self-contempt, in whimpering, feigned remorse and a merely passive attitude of mind – none of these have anything to do with that which can be counted as the beauty, let alone the sublimity, of a mentality. (AA V: 273)

This is a clear rejection of what I have called faculty-fixism, since it is hard not to recognize in these passages Rousseau's admonition that we remain in our limitations and do not seek to overcome them. For Kant, the goal is to overcome these limitations, even if they may ultimately be there, and "seek out the powers that still remain in us". We can recognize here the idea, presented in 6.2 as well, that we cannot tell in advance which are the powers of man, and that only a historical attempt to cultivate and develop our powers can reveal this.

Kant's theory of the sublime, then, resonates with his historical works in that the experience of the sublime opens to us the idea of man, not as what he is, but as what he could be, as Kant put it in his preface to the *Anthropology* from a pragmatic point of view:

A doctrine of the knowledge of the human being, systematically formulated (anthropology), can exist either in a physiological or in a pragmatic point of view. Physiological knowledge of the human being concerns the investigation of what *nature* makes of the human being; pragmatic, the investigation of what *he* as a free-acting being makes of himself, or can and should make of himself. (AA VII: 119)

And yet, Kant's discussion also indicates that there is an upper bound to the malleability of man: for the distinction between the lower and the higher faculties, and the fact that our cognition is discursive, will never be overcome. We can hope to expand and harmonize our capacities further and further, but we will never acquire a single intellectual nature, will never become so as to have direct conceptual contact with the particular. Since for Kant, it is the discursivity of our knowledge that marks us as finite, he believes that although we may be indefinitely malleable, we will never lose our finite nature. As Kant famously said: "out of such crooked wood as the human being is made, nothing entirely straight can be fabricated" (AA VIII: 23)

The whole idea of the vocation of man thus points towards the idea of epigenesis by resonating with the historical writings where, as we saw, an epigenesist picture of the nature of man was at stake. But there is of course another connection. In 6.2.3, I discussed Tetens' implicit connection of the perfectibility of man with the epigenesis of the faculties. That Kant was influenced by Tetens, that he spoke of epigenesis of faculties in the historical writings and that his *Critique of the Power of Judgment* both elaborates on the organic conception of the system of faculties and connects it with the idea of the vocation of man, all suggest that behind the Critical writings lay a genuine

idea of the epigenesis of the faculties. In the next section, I will try to extract such a theory.

## 7.4 The Epigenesis of the Faculties

At the close of the first part of this dissertation, I listed some features that a historicized transcendental idealism would need, and argued that these features matched up with the crucial features of Kant's conception of epigenesis identified in chapter 2. I then bracketed that issue in order to investigate Kant's conception of systematicity in part 2. I closed that part with the question of whether the system of faculties can be considered to be epigenetic in the same way that organisms can be. In the previous subsection, I have argued that this may very well be so. Now, we can put the pieces together and see what the epigenetic nature of the faculties might entail.

In the first place, we can see how Reason is the capacity for historicity as such: it confronts us with the limitations of our own capacities in such a way as to prompt us to overcome these limitations. In Kant's theory, it is through the experience of the sublime that the *Gemüt* is so prompted. Reason thus does not appear here as a transhistorical given, but rather as the ground for us to enter into history, to change, to become dynamic.

The third *Critique* thereby also identifies the productive or epigenetic power. Contrary to what one might expect, this power is not reason itself, but judgment. It is through judgment that the faculties are related to each other, and it is in judgment that the general function of the system of faculties is realized, namely cognition. In order to be able to cognize properly, the subjective condition for judgment must obtain, and this condition is the proper alignment of the faculties.

In chapter 4, we saw that, according to Kant, an organism is properly understood as a functional, normative unity of diverse interproducing and interspecifying parts. Its unity does not consist in a transcendent principle, but rather in a norm governing the unity of the organism. An organism is to be understood as acting in accordance with this norm, i.e. as striving to instantiate it. The third *Critique* revealed to us that there such a norm equally governs the unity of the faculties, and that this unity is not guaranteed by a leading or transcendent principle. It is in judging that it is revealed whether this norm obtains.

We have seen in this chapter that the norm does not automatically obtain, and that is contingent whether it obtains. The fact that we feel aesthetic pleasure or displeasure at all is, according to Kant, a sign that the different faculties are not eternally adjusted to each other, and the unity of the faculties can be "threatened" by some experiences. In

fact, in the sublime, we experience such a threat to that unity, since imagination is revealed as incapable of delivering or presenting something that can be grasped by conceptual means.

In chapter 3, I indicated that a historicized transcendental idealism needs to show how the transcendental is itself open to contingency. Now we see a way in which it is so, since the system of faculties itself is open to contingency in the described way. It is important to see, however, that its openness to contingency is not direct, but rather internal. It is not really so that aesthetic judgment reveals the mutual adaptedness between an object and our system of faculties. In fact, the object does not really figure in this judgment at all, although it is properly regarded as the contingent occasion for that judgment. What is judged is the mutual adaptedness of our faculties, i.e. the internal interadaptedness of our epistemic capacities, of our system of cognition.

The internal ground of the system of faculties' openness to cognition is thus the distinctness of the faculties, most importantly their division between intuitive and disjunctive. According to Kant, we are necessarily divided between the capacity for particularity and the capacity for generality. These two capacities are always only contingently in harmony – only in an intuitive intellect would they necessarily harmonize, precisely because they are not distinct at all.

This allows us to see the reason for what many have taken to be Kant's concessions to preformationism, namely his doctrine of germs and dispositions and his concomitant denial of a single generative force. In chapter 2, I argued that Kant introduced these concepts not to insist on preformation, but rather to suggest that epigenesis requires constraints. In chapter 3, I argued that Kant regarded the fixed germs and predispositions as those capacities which are required for historicity and epigenesis to be possible at all. In the current context, it is revealed that the distinctness of sensibility and understanding, or of lower and higher faculties, is such a constraint. Kant wants to deny that sensibility and understanding can originate from a same stem, not because he believes all capacities must be fully present at the onset, but rather because it is the division between the two that makes the *Gemüt* open to contingency at all. If there were not at least some distinction or articulation in the initial phases of an epigenetic process, then no such process would be prompted, since there would be no openness to contingency, and for Kant the contingent is required to prompt epigenetic processes.

Kant's insistence on the impossibility to ever overcome this distinction also indicates that we will never enter into a state where no further change is possible. That is to say, there can be no end of history, after which our system of faculties will no longer meaningfully change. Such an end of history is possible only if we fully overcome the division, and acquire an intuitive intellect, which is an impossibility for us. This reveals a major difference between Kant and Hegel: for Hegel, the intuitive intellect is possible, and hence there is a determinate teleological end-point to history. For Kant, such an

end-point will for ever remain impossible, and hence we cannot even in advance know in which direction we are going.

We can also better understand how exactly the contingent occasions the epigenesis of the faculties in Kant through the contrast with Hegel. For Hegel, there is a specific internal contradiction in each phase, and the contingent merely provides the occasion for this internal contradiction to come to the fore. For Kant, there need be no single or determinate tension that is then triggered by the contingent. What specific change is prompted is determined by the structure of the contingency triggering it. This contrasts with Hegel, for whom it is the specific change that can be prompted that determines what kind of contingency can trigger it. Moreover, for Kant there will usually not be change unless there are many tensions operating, up to a point where a readjustment or realignment of the faculties.

We must now, however, try to understand the nature of that change. Like Tetens, Kant seems to have taken the epigenetic process to be prompted by the relations between the (more) fundamental faculties rather than as an internal self-development of a single faculty. Like with Tetens, this does not need to mean that these fundamental faculties are entirely fixed or that no new faculties are possible. In fact, in Tetens' "epigenesis through evolution", the mutual interaction of the faculties can change the existing faculties and create new ones that are dependent upon their interactions. We can thus take quite seriously Kant's claim that we do not know what faculties we can have, and how far our faculties can be perfected. Moreover, we can see why the existing faculties are the predispositions to future ones: it is because of their structure and interrelations that future faculties are possible at all.

I believe we can also understand how there can be an epigenesis of the a priori in Kant's system if we keep in mind his organic conception. As we saw in 6.4, Kant seems to have regarded the a priori as a system of organically interadapted principles – interadapted to the general aim of cognition. This is because the different a priori principles are the principles of the different faculties. Now, if it is true that the a priori principles are to be understood through the other principles and the relations obtaining between them, then the principles can change if these relations change. In altering our faculties, their principles may ultimately undergo change as well.

It is important to realize, however, that the principles, such as the intuitions of space and time and the categories, can never be directly at stake in an adjustment, since within a given configuration they cannot fail. What may happen, however, is that concrete judging on the empirical and the contingent may ultimately provoke such changes in the system of faculties that the a priori itself becomes altered. It is thus only in the third degree that the a priori is open to contingency – but this does not mean that it cannot be open to contingency at all.

## 7.5 Conclusion

In this chapter, I have attempted to show how Kant may have thought of the epigenesis of reason by discussing the way in which he articulated his organic conception of the system of faculties in the *Critique of the Power of Judgment*. I have shown that he there conceives of the *Gemüt* as fundamentally open to contingency, and that he thereby tightens the analogy between it and an organism. I have also tried to show how the sublime can prompt change of the faculties, and open the *Gemüt* to historicity. All this is meant to show how the analyses of the previous chapters, mostly chapter 3 and chapter 6, converge upon a picture of Kantian epigenesis of reason as the historicity of the a priori.

In this dissertation, I have attempted to argue that Kant's epigenesis of reason could have implied more historicity than is traditionally assumed. I have done this by arguing that, according to Kant, the system of faculties is, like an organism, a plastic, i.e. epigenetic normative unity of a diversity of distinct and irreducible powers, and that the a priori depends on this interrelation. I will not here return to the many arguments of the dissertation, but instead indicate some future lines of research, and then briefly comment on the nature and scope of my claim.

In the first place, the present dissertation calls for a reappraisal of Kant's system of a priori principles. We need to understand two things, namely 1) how they are dependent upon the nature of the faculty to which they belong, and 2) how they mutually specify each other in the way suggested by Kant. If we can answer these questions, we can properly assess how other systems of a priori principles are possible.

Secondly, this dissertation suggests that reflective judgment plays a much more central role in Kant's thought than has traditionally been assumed. We should thus reappraise this theory. For instance, judgment seems to be peculiar in Kant's system because it has no principle. We may wonder what this means for epistemology and logic as conceived within the transcendental tradition.

Thirdly, there is the question of transcendental idealism. Throughout this dissertation, I have avoided dealing with the doctrine of transcendental idealism directly. I have, however, given several indications of how this doctrine should be understood. For instance, I have invariably refused to characterize it as a metaphysical position, and presented as a position *on* metaphysics, namely as impossible. I have indicated instead that I think of transcendental idealism as the theory that obligation, ought, or the deontological is more fundamental than causation, is, or the ontological. Additionally, I have suggested that any transcendental idealist needs to accept that the particular and the general need to be regarded as fundamentally distinct. What this entails for a complete picture of transcendental idealism, remains to be seen.

In this dissertation, however, I have attempt to argue that transcendental idealism should in any case not be conflated with Platonic Idealism. According to platonic idealism, the a priori and that from which the epistemic derives its normative thrust are eternal and unchanging. On more metaphysical versions, it sees it as eternal and unchanging because it is situated in some distinct ontological realm, be it the Fregean *dritte Reich* or the Leibnizian divine mind. I have tried to defend the idea that, for Kant, the transcendently ideal is more open to contingency than is traditionally assumed. In this, I have tried to vindicate him, and transcendental idealism in general, from the criticisms of 20<sup>th</sup> century philosophers who believe that everything that is not straightforwardly natural is supernatural. In chapters 4 and 5, there are hints of an interpretation of Kant that reveals this inference to be a non-sequitur. These hints, of course, await articulation. For now, I can dismiss as a *petitio principii* the most obvious criticism that may be raised against my reading, namely that it conflates the transcendental and the psychological. My response to this is simple: it assumes precisely what needs to be proven, namely that transcendental idealism is akin to platonic idealism in that it insists that the ideal is not open to contingency, which means that wherever we are concerned with contingency, we must be speaking of the psychological or of the messy concrete empirical interactions of natural entities. In this dissertation, I have argued that Kant sought to escape this dualism. One may ultimately come to the conclusion that no such third position can be coherent. But what one may not do is dismiss it before assessing it, and insist that those who arguably hold it cannot conceivably hold it.

There is of course another straightforward question that one may ask concerning the picture of Kantian epigenesis I offered in 7.4, namely whether Kant really held it. This, however, is a wrong question. Kant himself stated that the originator of a science or a system does not yet see all that it contains or entails, and that it is up to later thinkers to articulate what lies in its purview. All that I have wanted to show is that this too lies within the purview of Kant's transcendental idealism. Like all Kantians, I intend to go beyond Kant in a direction somehow indicated by Kant himself. I have simply tried to uncover one such direction: a radical perspective allowing for revolutions in the a priori. This Jacobin thrust of Kant's philosophy, however, has in the past centuries been expertly neutralized by the Thermidorian attempt to separate the fickleness of the natural from the stability of the transcendent, and the many ploys to restore, in however altered a form, the *ancien régime* of metaphysics.





# Bibliography

- Adams, R. M. (1994). *Leibniz: Determinist, Theist, Idealist*. Oxford: Oxford University Press.
- Adickes, E. (1924). *Kant als Naturforscher*. Berlin: Walter de Gruyter.
- Ahnert, T. "Newtonianism in Early Enlightenment Germany, c. 1720 to 1750: Metaphysics and the Critique of Dogmatic Philosophy", *Studies in the History and Philosophy of Science* 35: 471-491.
- Alighieri, D. (1904). *De Monarchia* (ed. & trans. Henry, A.). Boston / New York: Houghton, Mifflin & Co.
- Allison, H. E. (1996). *Idealism and Freedom: Essays on Kant's Theoretical and Practical Philosophy*. Cambridge: Cambridge University Press.
- Allison, H. E. (2001). *Kant's Theory of Taste*. Cambridge: Cambridge University Press.
- Allison, H. E. (2004). *Kant's Transcendental Idealism: An Interpretation and Defense* (2nd, revised edition). New Haven: Yale University Press.
- Ameriks, K. (2009). "The Purposive Development of Human Capacities," in Rorty, A. O. & Schmidt, J. (eds.) *Kant's Idea for a Universal History with a Cosmopolitan Aim: A Critical Guide*. Cambridge: Cambridge University Press: 46-67.
- Anderson, F. H. (1948). *The Philosophy of Francis Bacon*. Chicago: The University of Chicago Press.
- Aquinas, T. (1954). *In octo libros physicorum Aristotelis expositio* (ed. Maggiolo, P. M.). Turin/Rome: Marietti.
- Aquinas, T. (2014). *Basic Works* (ed. Hulse, J. & Pasnau, R.). Indianapolis: Hackett.
- Ariew, R. (1992). "Descartes and the Tree of Knowledge", *Synthese* 92(1): 101-116.
- Aristotle (1960). *Posterior Analytics; Topica* (ed. & trans. Forster, E. F.). Cambridge, Mass.: Harvard University Press.
- Arnauld, A. & Nicole, P. (1992). *La logique ou l'art de penser*. Paris: Gallimard.
- Bacon, F. (1875). *The Works of Francis Bacon* (ed. Ellis, R. E., Heath, D. D. & Spedding, J.) (Vol. 4). London: Longmans & co.; Simpkin, Marshall & Co.; Hamilton & co; Whittaker & co.; J. Bain.; Hodgson & co.; Richardson & co.; Houlston & sons; Bickers & son; H. Sotheran & co.; J. Cornish & Sons; Snow & co.; Aylott & son; and Hall & co.
- Bacon, F. (1876). *The Works of Francis Bacon* (ed. Ellis, R. E., Heath, D. D. & Spedding, J.) (Vol. 3). London: Longmans & co.; Simpkin, Marshall & Co.; Hamilton & co; Whittaker & co.; J. Bain.; Hodgson & co.; Richardson & co.; Houlston & sons; Bickers & son; H. Sotheran & co.; J. Cornish & Sons; Snow & co.; Aylott & son; and Hall & co.
- Bacon, F. (1878). *The Works of Francis Bacon* (ed. Ellis, R. E., Heath, D. D. & Spedding, J.) (Vol. 6). London: Longmans & co.; Simpkin, Marshall & Co.; Hamilton & co; Whittaker & co.; J. Bain.; Hodgson & co.; Richardson & co.; Houlston & sons; Bickers & son; H. Sotheran & co.; J. Cornish & Sons; Snow & co.; Aylott & son; and Hall & co.
- Balme, D. (1987). "Teleology and Necessity", in Gotthelf, A. & Lennox, J. G. (eds.). *Philosophical Issues in Aristotle's Biology*. Cambridge: Cambridge University Press: 275-285.
- Baumgarten, A. G. (1750). *Aesthetica*. Frankfurt a. d. Oder: Kleyb.

- Baumgarten, A. G. (1779). *Metaphysica*. Haller: Hemmerde.
- Beiser, F. (2009). *Diotima's Children: German Aesthetic Rationalism from Leibniz to Lessing*. Oxford: Oxford University Press.
- Bennett, J. (1974). *Kant's Dialectic*. Cambridge: Cambridge University Press.
- Berkeley, G. (2009). *Philosophical Writings* (ed. Clarke, D. M.). Cambridge: Cambridge University Press.
- Biener, Z. & Smeenk, C. (2012). "Cotes' Queries: Newton's Empiricism and Conceptions of Matter", in Janiak, A. & Schliesser, E. (eds.). *Interpreting Newton: Critical Essays*. Cambridge: Cambridge University Press: 105-137.
- Blumenbach, J. F. (1817). *The Institutions of Physiology* (trans. Elliotson, J.). London: Bensley & son.
- Blumenbach, J. F. (1789). *Über den Bildungstrieb*. Göttingen: Johann Christian Dieterich.
- Bonnet, C. (1779). *Oeuvres d'histoire naturelle et de philosophie de Charles Bonnet* (Vol. 3). Neuchatel: Samuel Fauche.
- Bono, J. J. (1990). "Reform and the Languages of Renaissance Theoretical Medicine: Harvey versus Fernel", *Journal of the History of Biology* 23(3): 341-387.
- Boscovich, R. (1745). *De Viribus Vivis*. Rome: Venantii Monaldini Bibliopolae.
- Bowler, P. J. (1971). "Preformation and Pre-existence in the Seventeenth Century: A Brief Analysis," *Journal of the History of Biology* 4(2): 221-244.
- Bowler, P. J. (1971). "Bonnet and Buffon: Theories of Generation and the Problem of Species," *Journal of the History of Biology* 6(2): 259-281.
- Boyle, R. (1725). *The Philosophical Works of the Honourable Robert Boyle, Esq.* (ed. Shaw, P.) (Vol. 2). London: Innys; Osborn & Longman.
- Brandom, R. B. (1994). *Making it Explicit: Reasoning, Representing and Discursive Commitment*. Cambridge, Mass.: Harvard University Press.
- Breidert, W. (2007). "Leonhard Euler and Philosophy", in Bradley, R. E. & Sandifer, C. E. (eds.), *Leonhard Euler: Life, Work and Legacy*. Amsterdam, Elsevier: 97-108.
- Breitenbach, A. (2008). "Two Views on Nature: A Solution to Kant's Antinomy of Mechanism and Teleology", *British Journal for the History of Philosophy* 16(2): 351-369.
- Brittan, G. (1978). *Kant's Theory of Science*. Princeton: Princeton University Press.
- Brittan, G. (1986). "Kant's Two grand Hypotheses", in Butts, R. E. (ed.), *Kant's Philosophy of Physical Science. Metaphysische Anfängsgründe der Naturwissenschaft 1786-1986*. Dordrecht: Reidel: 61-94.
- Broadie, S. (2011). *Nature and Divinity in Plato's Timaeus*. Cambridge: Cambridge University Press.
- Brook, A. (1994). *Kant and the Mind*. Cambridge: Cambridge University Press.
- Buchdahl, G. (1969). *Metaphysics and the Philosophy of Science: the Classical Origins from Descartes to Kant*. Oxford: Blackwell.
- Buffon, G.-L. (1749a). *Histoire naturelle, générale et particulière, avec la description du cabinet du roi* (Vol. 1). Paris: Imprimerie Royale.
- Buffon, G.-L. (1749b). *Histoire naturelle, générale et particulière, avec la description du cabinet du roi* (Vol. 2). Paris: Imprimerie Royale.
- Burke, E. (1764). *A Philosophicall Enquiry into the Origin of our Ideas of the Sublime and the Beautiful* (4<sup>th</sup> ed.). London: Dodsley.
- Butts, R. E. (1986). "Introduction: Kant's Quest for a Method for Metaphysics", in Butts, R. E. (ed.), *Kant's Philosophy of Physical Science. Metaphysische Anfängsgründe der Naturwissenschaft 1786-1986*. Dordrecht: Reidel: 1-22.
- Canguilhem (1991). *The Normal and the Pathological* (trans. Fawcett, C. R.). New York: Zone Books
- Canguilhem, G. (2008). *Knowledge of Life*. New York: Fordham University Press.
- Cantor, G. (1982). "Ueber eine elementare Frage der Mannigfaltigkeitslehre", in Meyer, W. F. (ed.). *Jahresbericht der deutschen Mathematiker-vereinigung* (Vol. 1). Berlin: Georg Reimer :75-78
- Cartwright, N. (1994). *Natures Capacities and Their Measurements*. Oxford: Oxford University Press.

- Cassirer, E. (1902). *Leibniz' System in seiner wissenschaftlichen Grundlagen*. Marburg: N. G. Elwert'sche Verlagsbuchhandlung.
- Cassirer, E. (1922). *Das Erkenntnisproblem in der Philosophie und Wissenschaft der neueren Zeit Band 1*. Berlin: Verlag Bruno Cassirer.
- Chang, H. (2009). "We Have Never been Wiggish (About Phlogiston)", *Centaurus* 51(4): 239-264.
- Chisholm, R. M. (1946). "The Contrary-to-Fact Conditional", *Mind* 55(220): 289-307.
- Cohen, A. (2009). *Kant and the Human Sciences: Biology, Anthropology and History*. Basingstoke: Palgrave Macmillan.
- Cohen, H. (1883). *Von Kants Einfluß auf die deutsche Kultur*. Berlin: Ferdinand Dümmler Verlag.
- Cohen, H. (1885). *Kants Theorie der Erfahrung* (2<sup>nd</sup> ed.). Berlin: Ferdinand Dümmler Verlag.
- Cohen, I. B. (1999). "A Guide to Newton's Principia", in Cohen, I. B. & Whitman, A. (eds.), *Isaac Newton's Principia: Mathematical Principles of Natural Philosophy*. Berkeley: University of California Press: 1-370.
- Cole, F. J. (1930). *Early Theories of Sexual Generation*. Oxford: Oxford University Press.
- Cooper, J. M. (1987). "Hypothetical Necessity and Natural Teleology", in Gotthelf, A. & Lennox, J. G. (eds.), *Philosophical Issues in Aristotle's Biology*. Cambridge: Cambridge University Press: 243-274.
- Costabel, P. (1973). *Leibniz and Dynamics: The texts of 1692* (trans. Maddison, R. E. W.). London: Methuen.
- Couturat, L. (1901). *La logique de Leibniz, d'après des documents inédits*. Paris: Felix Alcan.
- Crusius, C. A. (1766). *Entwurf der nothwendigen Vernunftwahrheiten wiefern sie den zufälligen entgegengesetzt werden*. Leipzig: Gleditschen.
- Cunningham, A. (1991). "How the "Principia" Got Its Name: Or, Taking Natural Philosophy Seriously", *History of Science* 29(4): 377-392.
- D'Alembert, J. le Rond (1758). *Traité de Dynamique* (2nd ed). Paris: Librairie David.
- D'Alembert, J. le Rond (1821). *Œuvres Complètes* (Vol. 1) Paris: Belin; Bossange père et fils; Bossange frères.
- Davidson, D. (2001). "On the Very Idea of a Conceptual Scheme", in *Inquiries into Truth and Interpretation*. Oxford: Oxford University Press: 183-198.
- Demarest, B. (2013a). "From the More Geometrico to the More Algebraico: d'Alembert and the Enlightenment's Transformation of Systematic Order". *Philosophica* 88: 71-102.
- Demarest, B. (2013b). "Novum sub Sole: Organicity and Temporality in Kant and Bergson", *International Journal of Computing Anticipatory Systems* (In press).
- Derrida, J. (1989). *Edmund Husserl's Origin of Geometry: An Introduction* (ed. & trans. Leavey, J. P.). Lincoln: University of Nebraska Press.
- Descartes, R. (1985). *The Philosophical Writings of René Descartes Vol. 1* (ed. & trans. Cottingham, J., Stoothoff, R. & Murdoch, D.). Cambridge: Cambridge University Press.
- Des Chene, D. (1996). *Physiologia: Natural Philosophy in Late Aristotelian and Cartesian Thought*. Ithaca: Cornell University Press.
- Des Chene, D. (2000). *Life's Form: Late Aristotelian Conceptions of the Soul*. Ithaca: Cornell University Press.
- Des Chene, D. (2001). *Spirits and Clocks: Machine and Organism in Descartes*. Ithaca: Cornell University Press.
- Detlefsen, K. (2003). "Supernaturalism, Occasionalism, and Preformation in Malebranche", *Perspectives on Science* 11(4): 443-483.
- Detlefsen, K. (2006). "Explanation and Demonstration in the Haller-Wolff Debate", in Smith, J. E. H. (ed.). *The Problem of Animal Generation in Early Modern Philosophy*. Cambridge: Cambridge University Press: 235-261.
- Diderot, D. (1994). *Oeuvres Tome I: Philosophie*. (ed. Versini, L.). Paris: Robert Laffont.
- Downing, L. (2012). "Maupertuis on Attraction as an Inherent Property of Matter", in Janiak, A. & Schliesser, E. (eds.). *Interpreting Newton: Critical Essays*. Cambridge: Cambridge University Press: 280-298.

- Dretske, F. I. (1977). "Laws of Nature", *Philosophy of Science* 44(2): 248-268.
- Dretske, F. E. (1981). *Knowledge and the Flow of Information*. Cambridge, Mass.: The MIT Press.
- Earman, J., & Roberts, J. T. (2005). "Contact with the Nomic: A Challenge for Deniers of Humean Supervenience about Laws of Nature Part I: Humean Supervenience", *Philosophy and Phenomenological Research* 71(1): 1-22.
- Evans, G. (1982). *The Varieties of Reference* (ed. McDowell, J.). Oxford: Oxford University Press.
- Euler, L. (1748). "Reflexions sur l'espace et le tems", *Memoires de l'academie royale des sciences et belles lettres* IV: 324-333.
- Euler, L. (1842). *Lettres de L. Euler à une princesse d'Allemagne sur divers sujets de physique et de philosophie* (Vol. 1). Paris: Hachette.
- Euler, L. (1752). "Recherches sur l'origine des forces", *Histoire de l'Académie Royale des Sciences et des Belles-Lettres de Berlin* VI: 419-447.
- Fechner, G. (1860). *Elemente der Psychophysik* (Vol. 1). Leipzig: Breitkopf und Härtel.
- Fisher, R. A. (1918). "The Correlation between Relatives on the Supposition of Mendelian Inheritance", *Transactions of the Royal Society of Edinburgh* 52: 399-433.
- Fisher, R. A. (1930). *The Genetical Theory of Natural Selection*. Oxford: Clarendon Press.
- Fletcher, A. (2005). "Francis Bacon's Forms and the Logic of Ramist Conversion," *Journal of the History of Philosophy* 43(2): 157-169.
- Frankfurt, H. G. (2008). *Demons, Dreamers, and Madmen: The Defense of Reason in Descartes's Meditations*. Princeton: Princeton University Press.
- Friedman, M. (1986). "The Metaphysical Foundations of Newtonian Science", in Butts, R. E. (ed.), *Kant's Philosophy of Physical Science. Metaphysische Anfängsgründe der Naturwissenschaft 1786-1986*. Dordrecht: Reidel: 25-60.
- Friedman, M. (1992a). "Causal Laws and the Foundations of Natural Science", in Guyer, P. (ed.), *The Cambridge Companion to Kant*. Cambridge: Cambridge University Press: 161-199.
- Friedman, M. (1992b). *Kant and the Exact Sciences*. Cambridge, Mass.: Harvard University Press.
- Gabbey, A. (1980). "Force and Inertia in the Seventeenth Century: Descartes to Newton", in Gaukroger, S. (ed.), *Descartes: Philosophy, Mathematics, Physics*. Sussex: The Harvester Press: 231-320.
- Galen. (1916). *On the Natural Faculties* (ed. & trans. Brock, A. J.). Cambridge, Mass.: Harvard University Press.
- Garber, D. (1992). *Descartes' Metaphysical Physics*. Chicago: University of Chicago Press.
- Garber, D. (2000). "A Different Descartes: Descartes and the Programme for Mathematical Physics in his Correspondence", in Gaukroger, S., Schuster, J. & Sutton, J. (eds.). *Descartes' Natural Philosophy*. London: Routledge: 113-130.
- Garber, D. (2009). *Leibniz: Body, Substance, Monad*. Oxford: Oxford University Press.
- Garrett, D. (1999). "Teleology in Spinoza and Early Modern Rationalism", in Gennaro, R. J. & Huenemann, C. (eds.) *New Essays on the Rationalists*. Oxford: Oxford University Press: 310-335.
- Gassendi, P. (1727). *Opera Omnia I*. Firenze: Giovanni Tartini.
- Geiger, I. (2003). "Is the Assumption of a Systematic Whole of Empirical Concepts a Necessary Condition of Knowledge?", *Kant-Studien* 94: 273-298.
- Genova, A. C. (1974). "Kant's Epigenesis of Pure Reason," *Kant-Studien* 65: 259-273.
- Gigliani, G. (1997). "The machines of the body and the operations of the soul Marcello Malpighi's anatomy", in Bertoloni Meli, D. (ed.). *Marcello Malpighi: Anatomist and Physician*. Firenze: Olschki: 149-174.
- Ginsborg, H. (2004). "Two Kinds of Mechanical Inexplicability in Kant and Aristotle", *Journal of the History of Philosophy* 42 (1): 33-65.
- Goldberg, B. (2013). "A dark business, full of shadows: Analogy and theology in William Harvey", *Studies in History and Philosophy of Biological and Biomedical Sciences* 44: 419-432.
- Goodman, N. (1947). "The Problem of Counterfactual Conditionals", *The Journal of Philosophy* 44(5): 113-128.

- Goodman, N. (1977). *The Structure of Appearance*. Dordrecht: Springer.
- Gould, S. J. (1977). *Ontogeny and Phylogeny*. Cambridge, Mass.: Harvard University Press.
- Gould, S. J. (1979). "The spandrels of San Marco and the Panglossian Paradigm: a critique of the adaptationist programme", *Proceedings of the Royal Society London B* 205: 581-598.
- Gould, S. J. (2002). *The Structure of Evolutionary Theory*. Cambridge, Mass.: Harvard University Press.
- Gregory, A. (2001). "Harvey, Aristotle and the Weather Cycle". *Stud. Hist. Phil. Biol. & Biomed. Sci* 32(1): 153-168.
- Grier, M. (2001). *Kant's Doctrine of Transcendental Illusion*. Cambridge: Cambridge University Press.
- Griffiths, P. E. & Gray, R. D. (1994). "Developmental Systems and Evolutionary Explanation", *The Journal of Philosophy* 91(6): 277-304.
- Gueroult, M. (1967). *Leibniz. Dynamique et Metaphysique*. Paris: Aubier - Montaigne.
- Guyer, P. (1987). *Kant and the Claims of Knowledge*. Cambridge: Cambridge University Press.
- Guyer, P. (1997). *Kant and the Claims of Taste* (2<sup>nd</sup> ed.). Cambridge: Cambridge University Press.
- Haldane, J. B. S. (1924). "A Mathematical Theory of Natural and Artificial Selection Part I", *Transactions of the Cambridge Philosophical Society* 23: 19-41.
- Hall, T. S. (1970). "Descartes' Physiological Method: Position, Principles, Examples", *Journal of the History of Biology* 3(1): 53-79.
- Haller, A. von (1758). *Sur la formation du coeur dans le poulet* (Vol. 2). Lausanne: Bousquet.
- Hamilton, W. R. (1837). "Theory of Conjugate Functions, or Algebraic Couples; with a Preliminary and Elementary Essay on Algebra as the Science of Pure Time", *Transactions of the Royal Irish Academy* 17(1): 293-422.
- Hankins, T. L. (1965). "Eighteenth-Century Attempts to Resolve the Vis viva Controversy", *Isis* 56(3): 281-297.
- Harman, P. M. (1983). "Force and Inertia: Euler and Kant's *Metaphysical Foundations of Natural Science*", In Shea, W. R. (ed.), *Nature Mathematized: Historical and Philosophical Case Studies in Classical Modern Philosophy*. Dordrecht: Reidel: 229-249.
- Harvey, W. (1847). *The Works of William Harvey* (ed. & trans. Willis, R.). London: Sydenham Society.
- Hegel, G. W. F. (1977). *Faith and Knowledge* (trans. Cerf, W. & Harris, H. S.). Albany: State University of New York Press.
- Hegel, G. W. F. (1991). *The encyclopaedia logic, with the Zusätze: Part I of the Encyclopaedia of philosophical sciences with the Zusätze* (ed. & trans. Geraets, T. F., Suchting, W. A. & Harris, H. S.). Indianapolis: Hackett.
- Heidegger, M. (1962). *Kant and the Problem of Metaphysics* (trans. Churchill, J. S.). Bloomington: Indiana University Press.
- Heidegger, M. (2000). *Introduction to Metaphysics* (trans. Fried, G. & Polt, R.). New Haven: Yale University Press.
- Heine, H. (1861). *Sämmtliche Werke* (Vol. 5.1). Hamburg: Hoffman und Campe.
- Henrich, D. (1994). *The Unity of Reason: Essay's on Kant's Philosophy* (ed. Velkley, R. L.; trans. Edwards, J., Hunt, L., Kuehn, M. & Zoeller, G.). Cambridge, Mass.: Harvard University Press.
- Herder, J. G. (1869). *Ideen zur Geschichte der Menschheit* (Vol. 1). Leipzig: Brockhaus.
- Hess, H.-J. (1978). "Die unveröffentlichten naturwissenschaftlichen und technischen Arbeiten von G.W. Leibniz aus der Zeit seines Parisaufenthaltes. Eine Kurzcharakteristik", *Studia Leibniziana Supplementa* 17(1): 183-217.
- Hessbrüggen-Walter, S. (2004). *Die Seele und ihre Vermögen : Kants Metaphysik des Mentalen in der "Kritik der reinen Vernunft"*. Paderborn: Mentis.
- Hippocrates. (2012). *Works* (Vol. 10) (ed. & trans. Potter, P.). Cambridge, Mass.: Harvard University Press.
- Hobbes, T. (1839). *The English Works of Sir Thomas Hobbes* (ed. William Molesworth) (Vol. 1). London: John Bohn.

- Hooke, R. (1667). *Micrographia: or, Some physiological descriptions of minute bodies made by magnifying glasses*. London: James Allestry.
- Hume, D. (2009) *A Treatise of Human Nature* (ed. Norton, D. F. & Norton M. J.). Oxford: Oxford University Press.
- Huneman, P. (2008). *Métaphysique et biologie. Kant et la constitution du concept d'organisme*. Paris: Kimé.
- Hutcheson, F. (2004). *An Inquiry into the Original of Our Ideas of Beauty and Virtue in Two Treatises* (ed. Leidhold, W.). Indianapolis: Liberty Fund.
- Huygens, C. (1669). "A Summary Account of the Laws of Motion", *Philosophical Transactions* 4: 925-928.
- Iltis, C. (1970). "D'Alembert and the Vis Viva Controversy", *Studies in History and Philosophy of Science Part A* 1(2): 135-144.
- Iltis, C. (1971). "Leibniz and the Vis Viva Controversy", *Isis* 62(1): 21-35.
- Israel, J. (2001). *Radical Enlightenment: Philosophy and the Making of Modernity 1650-1750*. Oxford: Oxford University Press.
- Jack, J. (2009). "A Pedagogy of Sight: Microscopic Vision in Robert Hooke's *Micrographia*", *Quarterly Journal of Speech*: 95(2): 192-209.
- Janiak, A. (2008). *Newton as Philosopher*. Cambridge: Cambridge University Press.
- Jones, J.-E. (2005). "Boyle, Classification and the Workmanship of the Understanding Thesis". *Journal of the History of Philosophy*, 43(2): 171-183.
- Kemp Smith, N. (2003). *A Commentary to Kant's 'Critique of Pure Reason'*. Houndmills, Basingstoke, Hampshire: Palgrave Macmillan
- Kitcher, P. (1984). "Kant's Philosophy of Science", in Wood, A. W. (ed.). *Self and Nature in Kant's Philosophy*. Ithaca: Cornell University Press: 185-215.
- Kitcher, P. (1996). "Aprioristic Yearnings". *Erkenntnis* 44: 397-416.
- Kleingeld, P. (1999). "Kant, History and the Idea of Moral Development", *History of Philosophy Quarterly* 16(1): 59-80.
- Koyré, A. (1966). *Études Galiléennes*. Paris: Hermann.
- Kreines, J. (2005). "The Inexplicability of Kant's *Naturzweck*: Kant on Teleology, Explanation and Biology", *Archiv für Geschichte der Philosophie* 87: 270-311.
- Leeuwenhoek, A. van (1686). *Ontleding en Ontdekkingen van leevende Dierkens in de Teel-Deelen van verscheide Dieren, Vogelen en Vissen; Van het Hout met der selver menigvuldige vaaten; Van Hair, Vlees en Vis; Als mede van de grote menigte der Dierkens in de excrementen*. Leiden: Cornelis Boutestyn.
- Leibniz, G. W. (1989). *Philosophical Essays* (ed. & trans. Ariew, R. & Garber, D.). Indianapolis: Hackett.
- Leibniz, G. W. (1996). *New Essays on Human Understanding* (ed. & trans. Bennett, J. & Remnant, P.). Cambridge: Cambridge University Press.
- Lennox, J. G. (2006). "The Comparative Study of Animal Development: William Harvey's Aristotelianism", in Smith, J. E. H. (ed.). *The Problem of Animal Generation in Early Modern Philosophy*. Cambridge: Cambridge University Press: 21-46.
- Lenoir, T. (1982). *The Strategy of Life: Teleology and Mechanics in Nineteenth Century German Biology*.
- Lewis, C. I. (1918). *A Survey of Symbolic Logic*. Berkeley: University of California Press.
- Lewis, D. K. (1973). *Counterfactuals*. Oxford: Blackwell.
- Lewis, D. K. (1986). *Philosophical Papers II*. Oxford: Oxford University Press.
- Locke, J. (1975). *An Essay concerning Human Understanding* (ed. & trans. Nidditch, P. H.). Oxford: Clarendon.
- Lovejoy, A. O. (1959). "Kant and Evolution," in Glass, B., Temkin, O. & Straus, W. L. (eds.). *Forerunners of Darwin: 1745-1859*. Baltimore: The Johns Hopkins Press: 173-206.
- Lovejoy, A. O. (1966). *The Great Chain of Being: A Study of the History of an Idea*. Cambridge, Mass.: Harvard University Press.

- Lovejoy, A. O. (1967). "On Kant's Reply to Hume", in Gram, M. S. (ed.). *Kant: Disputed Questions*. Chicago: Quadrangle Books: 284-308.
- Maienschein, J. (2000). "Competing Epistemologies and Developmental Biology", in Creath, R. & Maienschein, J. (eds.). *Biology and Epistemology*. Cambridge: Cambridge University Press: 122-137.
- Maier, A. (1955). *Metaphysische Hintergründe der spätscholastischen Naturphilosophie*. Rome: Edizioni di storia e letteratura
- Makkreel, R. A. (1990). *Imagination and Interpretation in Kant: the Hermeneutical Import of the Critique of Judgment*. Chicago: University of Chicago Press.
- Malpighi, M. (1687). *Marcelli Malpighi Operum Tomus Secundus*. Leiden: Peter vander Aa.
- Malpighi, M. (1697). *Marcelli Malpighi Opera Posthuma*. London: Churchill.
- Mariotte, E. (1673). *Traite de la Percussion du Chocq des Corps*. Paris: Estienne Michalet.
- Mates, B. (1986). *The Philosophy of Leibniz: Metaphysics and Language*. Oxford: Oxford University Press.
- Matthews, M. R. (ed.) (1989). *The Scientific Background to Modern Philosophy*: Indianapolis: Hackett.
- Maupertuis, P.-L. Moreau de (1768a). *Oeuvres de Maupertuis (Vol. 1)*. Lyon: Jean-Marie Bruyset.
- Maupertuis, P.-L. Moreau de (1768b). *Oeuvres de Maupertuis (Vol. 2)*. Lyon: Jean-Marie Bruyset.
- Mayr, E. (1997). *This is Biology. The Science of the Living World*. Cambridge, Mass.: Harvard University Press.
- McDowell, J. (1994). *Mind and World*. Cambridge, Mass.: Harvard University Press.
- McFarland, J. D. (1970). *Kant's Concept of Teleology*. Edinburgh: Edinburgh University Press.
- McLaughlin, P. (1990). *Kant's Critique of Teleology in Biological Explanation: Antinomy and Teleology*. Lewiston/Queenston/Lampeter: The Edwin Mellen Press.
- McLaughlin, P. (2000). "Force, determination and impact", in Gaukroger, S., Schuster, J. & Sutton, J. (eds.). *Descartes' Natural Philosophy*. London: Routledge: 81-112.
- McLaughlin, P. (2007). "Kant on Heredity and Adaptation," in Müller-Wille, S & Rheinberger, H.-J. (eds.). *Heredity Produced. At the Crossroads of Biology, Politics and Culture, 1500-1870*. Cambridge, Mass: MIT Press: 277-292.
- McRae (1957). "The Unity of the Sciences: Bacon, Descartes and Leibniz," *Journal of the History of Ideas* 18(1): 27-48.
- Mendelssohn, M. (2008). *Metaphysische Schriften* (ed. Vogt, W.). Hamburg: Meiner.
- Mensch, J. (2013). *Kant's Organicism: Epigenesis and the Development of Critical Philosophy*. Chicago: University of Chicago Press.
- Monod, J. (1970). *Le hasard et la nécessité: Essai sur la philosophie naturelle de la biologie moderne*. Paris: Seuil.
- Müller-Sievers, H. (1997). *Self-Generation. Biology, Philosophy, and Literature Around 1800*. Stanford: Stanford University Press.
- Nachtomy, O. (2014). "Infinity and Life: The Role of Infinity in Leibniz's Theory of Living Beings", in Nachtomy, O. & Smith, J. *The Life Sciences in Early Modern Philosophy*: Oxford University Press: 9-28.
- Needham, J. (1959). *A History of Embryology* (2nd ed). Cambridge: Cambridge University Press.
- Neiman, S. (1997). *The Unity of Reason: Rereading Kant*. Oxford: Oxford University Press.
- Nelson, A. (1996). "Micro-Chaos and Idealization in Cartesian Physics", *Philosophical Studies* 77(2): 377-391.
- Newton, I. (1999). "Mathematical Principles of Natural Philosophy", in Cohen, I. B. & Whitman, A. (eds.), *Isaac Newton's Principia: Mathematical Principles of Natural Philosophy*. Berkeley: University of California Press: 371-944.
- Nietzsche, F. (2002). *Beyond Good and Evil: Prelude to a Philosophy of the Future*. Cambridge: Cambridge University Press.
- Oyama, S. (2000). *The Ontogeny of Information. Developmental Systems and Evolution* (2nd ed., Revised and Expanded). Durham, NC: Duke University Press.

- Pagel, W. (1967). *William Harvey's Biological Ideas: Selected Aspects and Historical Background*. Basel / New York: S. Karger.
- Parkinson, G. H. R. (1965). *Logic and Reality in Leibniz's Metaphysics*. Oxford: Clarendon Press.
- Philonenko, A. (1972). *L'oeuvre de Kant: la philosophie critique* (Vol. 2). Paris: Vrin.
- Pinto-Correia, C. (1997). *The Ovary of Eve: Egg and Sperm and Preformation*. Chicago: University of Chicago Press.
- Popkin, R. (2003). *The History of Skepticism from Savonarola to Bayle*. Oxford: Oxford University Press.
- Porphyry (2003). *Introduction* (ed. & trans. Barnes, J). Oxford: Oxford University Press.
- Pradeu, T. (2012). *The Limits of the Self: Immunology and Biological Identity* (trans. Vitanza, E.). Oxford: Oxford University Press.
- Press, J. S., & Tanur, J. M. (2001). *The Subjectivity of Scientists and the Bayesian Approach*. New York: Wiley.
- Reinhold, K. L. (1795). *Versuch einer neuen Theorie des menschlichen Vorstellungsvermögens*. Prague / Iena: Widtmann & Mauke.
- Rescher, N. (1981). *Leibniz's Metaphysics of Nature. A Series of Essays*. Dordrecht: Reidel.
- Rescher, N. (2000). *Kant and the Reach of Reason. Studies in Kant's Theory of Rational Systematization*. Cambridge: Cambridge University Press.
- Riehl, A. (1908). *Der philosophische Kritizismus: Geschichte und System* (2<sup>nd</sup> ed. Vol. 1). Leipzig: Wilhelm Engelmann.
- Robert, J. S. (2004). *Embryology, Epigenesis and Evolution: Taking Development Seriously*. Cambridge: Cambridge University Press.
- Roe, S. A. (1981). *Matter, Life and Generation: Eighteenth-Century Embryology and the Haller-Wolff Debate*. Cambridge: Cambridge University Press.
- Roger, J. (1963). *Les sciences de la vie dans la pensée française du XVIIIe siècle: La generation des animaux de Descartes à l'Encyclopédie*. Paris: Armand Colin.
- Rosenberg, A. (1993). "Hume and the philosophy of science", in Norton, D. F. (ed.). *The Cambridge Companion to Hume*. Cambridge: Cambridge University Press: 64-89.
- Rousseau, J.-J. (1974). *Emile or on Education* (trans. Allan Bloom). New York: Basic Books.
- Russell, B. (1992). *A Critical Exposition of the Philosophy of Leibniz*. London: Routledge.
- Rutherford, D. (1995). *Leibniz and the Rational Order of Nature*. Cambridge: Cambridge University Press.
- Sabra, A. I. (1981). *Theories of Light from Descartes to Newton*. Cambridge: Cambridge University Press.
- Schliesser, E. (2013). "On reading Newton as an Epicurean: Kant, Spinozism and the changes to the Principia", in *Studies in History and Philosophy of Science* 44: 416-428.
- Schmitt, C. B., & Webster, C. (1989). "Harvey and M. A. Severino: A Neglected Medical Relationship", in Webster, C. (ed.). *Reappraisals in Renaissance Thought*. London: Variorum Reprints: 49-75.
- Schönfeld, M. (2000). *The Philosophy of the Young Kant: the Precritical Period*. Oxford: Oxford University Press.
- Sellars, W. (1997). *Empiricism and the Philosophy of Mind* (ed. Rorty, R. & Brandom, R. B.). Cambridge, Mass.: Harvard University Press.
- Shank, J. B. (2008). *The Newton Wars and the Beginning of the French Enlightenment*. Chicago: Chicago University Press.
- Shay, W. (1986). "Filled with Wonder: Kant's Cosmological Essay, *The Universal Natural History and Theory of the Heavens*", in Butts, R. E. (ed.), *Kant's Philosophy of Physical Science. Metaphysische Anfangsgründe der Naturwissenschaft 1786-1986*. Dordrecht: Reidel: 95-124.
- Simondon, G. (1989). *L'Individuation psychique et collective à la lumière des notions de Forme, Information, Potentiel et Métastabilité*. Paris: Aubier
- Sleigh, R. C. (1990). *Leibniz and Arnauld. A Commentary on their Correspondence*. New Haven: Yale University Press.



- Sloan, P. 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 (2): 229-253.
- Smith, J. E. H. (2011). *Divine Machines: Leibniz and the Sciences of Life*. Princeton: Princeton University Press.
- Speiser, A. (1934). *Leonhard Euler und die Deutsche Philosophie*. Zürich: Orell Füssli Verlag.
- Strawson, P. F. (1966). *The Bounds of Sense: An Essay on Kant's Critique of Pure Reason*. London: Methuen.
- Swammerdam, J. (1737). *Bybel der Natuure* (vol. 1). Leiden: Isaak Severinus, Boudewyn vander Aa & Pieter vander Aa.
- Tetens, J. H. (1777). *Philosophische Versuche über die menschliche Natur und ihre Entwicklung* (2 vols.). Leipzig: Weidmanns Erben und Reich.
- Timmerding, H. E. (1918). "Kant und Euler", *Kant-Studien* 23: 18-64.
- Tonelli, G. (1959). *Elementi metodologici e metafisici in Kant dal 1745 al 1768*. Torino: Edizioni di Filosofia.
- Van Strien, M. (2014). "On the Origins and Foundations of Laplacian Determinism", *Studies in the History and Philosophy of Science* 45: 24-31.
- Voltaire (1782). *Éléments de la Philosophie de Newton*. Lausanne: Jules Henri Pott.
- Wallis, J. (1668). "A Summary Account of the General Laws of Motion", *Philosophical Transactions* 3: 864-866.
- Walsh, D. W. (1975). *Kant's Criticism of Metaphysics*. Edinburgh: Edinburgh University Press.
- Waterlow, S. (1982). *Nature, Change and Agency in Aristotle's Physics: A Philosophical Study*. Oxford: Clarendon Press.
- Watkins, E. (2005). *Kant and the Metaphysics of Causality*. Cambridge: Cambridge University Press.
- Watkins, E. (2010). "The System of Principles", in Guyer, P. (ed.) *The Cambridge Companion to Kant and the Critique of Pure Reason*. Cambridge: Cambridge University Press: 151-167.
- Westfall, R. S. (1971). *Force in Newton's Physics: the Science of Dynamics in the Seventeenth Century*. London: Macdonald and Co.
- Whitaker, V. K. (1970). "Bacon's Doctrine of Forms: A Study of Seventeenth-Century Eclecticism," *Hunting Library Quarterly* 33(3): 209-216.
- Williams, B. (2005). *Descartes: The Project of Pure Inquiry*. London: Routledge.
- Wilson, C. (1995). *The Invisible World: Early Modern Philosophy and the Invention of the Microscope*. Princeton: Princeton University Press.
- Windelband, W. (1924) *Präludien: Aufsätze und Reden zur Philosophie und ihre Geschichte* (Vol. 1). Tübingen: Mohr-Siebeck.
- Wolfe, C. T. (2011). "From substantial to functional vitalism and beyond: animas, organisms and attitudes", *Eidos* 14: 212-235.
- Wolff, C. (1738). *Psychologia Empirica, Methodo Scientifica Pertractata*. Frankfurt & Leipzig: Renger.
- Wolff, C. (1747). *Vernünfftige Gedanken von Gott, der Welt und der Seele des Menschen, auch allen Dingen überhaupt*. Haller: Rengerische Buchhandlung.
- Wolff, C. F. (1768). *Theorie von der Generation*. Berlin: Wilhelm Birnstiel.
- Wood, A. W. (2003). *Karl Marx* (2<sup>nd</sup> ed.). London: Routledge.
- Wood, A. W. (2006). "Kant's Philosophy of History", in Kleingeld, P. (ed.). *Toward Perpetual Peace and Other Writings on Politics, Peace, and History*. New Haven: Yale University Press: 243-262.
- Wren, C. (1668). "Lex Naturae de Collisione Corporum", *Philosophical Transactions* 3: 867.
- Wright, M. R. (2000). "Myth, Science and Reason in the *Timaeus*", in Wright, M. R. (ed.). *Reason and Necessity: Essays on Plato's Timaeus*. London & Swansea: Duckworth & the Classical Press of Wales: 1-22.
- Wright, S. (1932). "The Roles of Mutation, Inbreeding, Crossbreeding and Selection in Evolution", *Proceedings of the Sixth International Congress on Genetics* 1: 355-366.
- Yovel, Y. (1980). *Kant and the Philosophy of History*. Princeton: Princeton University Press.

- Zammito, J. H. (1992). *The Genesis of Kant's Critique of Judgment*. Chicago: The University of Chicago Press.
- Zammito, J. H. (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.
- Zuckert, R. (2007). *Kant on Beauty and Biology: An Interpretation of the Critique of Judgment*. Cambridge: Cambridge University Press.
- Zumbach, C. (1984). *The Transcendent Science: Kant's Conception of Biological Methodology*. The Hague: Martinus Nijhoff.

