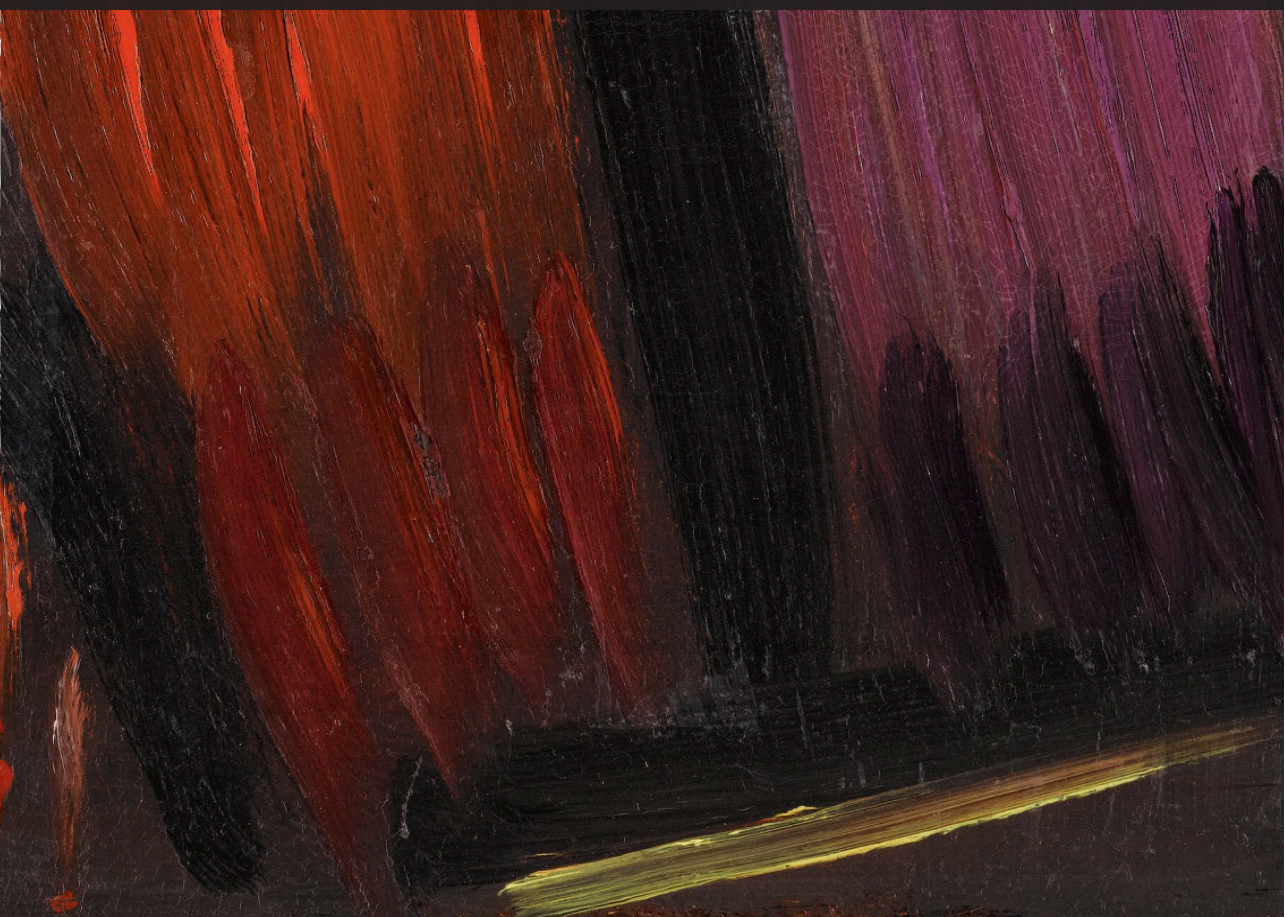
Conceptual Realism

The Structure of Metaphysical Thought

Jesse M. Mulder

θπ



Conceptual Realism

The Structure of Metaphysical Thought

Conceptueel Realisme

De structuur van metafysisch denken

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de rector magnificus, prof. dr. G.J. van der Zwaan, ingevolge het besluit van het college voor promoties in het openbaar te verdedigen op woensdag 18 juni 2014 des middags te 4.15 uur

door

Jesse Merijn Mulder

geboren op 28 oktober 1982

te Alkmaar

Promotoren: Prof. dr. A. Visser
Prof. dr. T. Müller

Painting on cover:
Alexej von Jawlensky, *Meditation* (unknown date).
Public domain, via Wikimedia Commons.

ISBN 978-94-6103-040-5
ISSN 0927-3395

Acknowledgments

Writing a dissertation is a solitary business, yet impossible without the dedicated support of many people, in many different ways. During the four years I have spent developing, revising, refining, and formulating the ideas to be found on the ensuing pages, I have benefited greatly from the able advice, friendly feedback, and constructive criticism that various people around me were kindly willing to offer.

The first person to whom I would like to express my gratitude is Thomas Müller, my (second) promotor. I was the first PhD-researcher of the rapidly growing group of PhDs and postdocs Thomas gathered around him through the research projects he managed to set up, and I have enjoyed working together with him in establishing and maintaining a fine, philosophically stimulating and socially pleasant research group. Thomas has followed my progress with caring attention, and, together with Albert Visser, my (first) promotor, he has let me find my own philosophical ways in great trust—I have been able to navigate my way through adventurous philosophical waters in the comforting certainty that both of them were there to help if need be.

I am especially thankful to my fellow PhD-researchers Antje Rumberg, Dawa Ometto and Niels van Miltenburg (in alphabetical order). I believe that our lively discussions, and all the time we spent reading each other's work and listening to each other's philosophical problems and puzzles, each other's frustrations and disappointments, and each other's enthusiastic discoveries and new ideas, have started to add a hint of seriousness to our running joke of starting our own 'Utrecht School of Philosophy'. They have made working on this thesis a great joy for me.

Although my philosophical upbringing took place in Utrecht—I enjoyed both my undergraduate and graduate studies at Utrecht University—the path that eventually led to my taking up metaphysics, the subject with which this dissertation deals, essentially involves an important detour through the University of Connecticut, and in particular involves two philosophers connected to that university, both of whom

I owe very much. The first is Ruth Millikan, the second Tim (Crawford) Elder. My first attempt at philosophical research, resulting in my bachelor's thesis, concerned Ruth's fascinating work on mind, language, and meaning. I was happy enough to be able to discuss the questions I had with Ruth herself, since she happened to be at Utrecht University for a couple of months at that time. I have found Ruth's lively and original way of doing philosophy, both in writing and in conversation, to be most inspiring. Two years later, in 2009, I visited UConn, and there I took a graduate course on metaphysics with Tim. His lively presentation of metaphysical questions and puzzles, as well as his well-justified discontent (at times even despair) with leading but bizarre metaphysical views, have fueled a lasting interest in metaphysics on my part. That interest found a first expression in my master's thesis on mereology and metaphysics, for which Tim served as second supervisor (next to Thomas), and culminated in the present essay.

I have found much inspiration for my own metaphysical views in the work of Michael Thompson, and so I was glad to have the opportunity to discuss my ideas with him during a short stay at Pittsburgh in early 2013. The influence of his original thought on my work especially shows in Part II of this thesis, which is built around his notion of 'conceptual gears'.

Many other people have contributed to improving this thesis, for instance with their feedback on either spoken or written versions of parts of its content, for which I would like to express my gratitude. First of all, the people in our research group at Utrecht: Daan Evers, Leszek Wroński, Michael De, and Rosja Mastop. Also other colleagues at our department: Herman Hendriks, Herman Philipse, Jan van Ophuijsen, Jan Vorstenbosch, Janneke van Lith, Joel Anderson, Joeri Witteveen, Joop Leo, Menno Lievers, Paul Ziche, Rosalie Iemhoff, and Vincent van Oostrom. The department of philosophy in Utrecht has provided me—and our research group in general—with a stimulating and friendly environment to conduct my research in. Furthermore, I would like to thank Alexander Reutlinger, Austen Clark, Bob Coenraats, Erik Schliesser, Fred Muller, Jan Willem Wieland, Jeroen de Ridder, Kit Fine, Lars Tump, Markus Schrenk, Marius Backmann, Michael Lynch, Rik Peels, Sebastian Rödl, Stefan Roski, and Verena Wagner.

I have had the fortunate opportunity to present material from several chapters of this dissertation at various conferences, workshops, and other occasions. The questions and remarks posed by attendees, and the subsequent discussions that often developed during breaks, have greatly helped me in developing key ideas of

this thesis. In particular, I would like to thank the attendants of the *Dutch Research Seminar in Analytic Philosophy* and of the *Theoretical Philosophy Lunch Lecture Series* at Utrecht University.

Above all, I thank my wife, Anne Geels. Anne has accompanied the entire four-year process culminating in this thesis with great patience and a healthy dose of humor. Living together with her and our two daughters Ronja and Momo, who were both born within the last two years, has been the best and most beautiful distraction I could have ever wished.

Jesse Mulder
Easter 2014

* * *

Contents

Introduction

Aristotle's *Metaphysics* opens with the famous statement: 'By nature, all men long to know' [Aristotle 1998, p. 4, 980a21]. And that is right: whether insofar as we are attempting to cope with the parochial surroundings of our own lives, as we all do, or insofar as we are contributing to research in some academic discipline or other, as only some of us do: we find it important to understand what we are dealing with, and to shape our lives in accordance with that understanding.

How do we go about finding our way in reality? Well, we observe, and we think. We pay attention to how our neighbor reacts to our morning greeting, and thereby learn to recognize what mood he is in today—so we know when to ask or rather not to ask him for help in the garden. We feel, as we drive in our car, how we are pulled against the left door a bit too much, and learn to decelerate somewhat more when negotiating that tricky curve. We observe what happens to bacteria in our petri dish after having added a bit of penicillin, and decide that that substance may well prove useful for medicinal purposes. We scrutinize the peculiar patterns in the light spectrum originating from a certain star, and conjecture that there is hydrogen present in the relevant remote regions of space.

In all such cases, we use our cognitive abilities to deepen our understanding of whatever it is we are dealing with—again, we observe, and we think. With observation alone, there would be just the 'blooming, buzzing confusion' of experience[1]: understanding is only possible when we start to think. If we take *concepts*, in a pre-theoretical, intuitive sense, to be the elements of thought, we may say that such understanding is *conceptual*.

Thus, in our examples, we find that we are dealing with the relevant bits of reality by employing the concepts *human being*, *character* and *mood*; *car*, *acceleration*, and *curve*; *bacteria* and *penicillin*; and *star*, *hydrogen*, and *light*—amongst many others. If

[1] The phrase comes from William James [1890, p. 462].

we may use such an unusual kind of ostension: think a while about each of our examples, about the way in which the relevant concepts are brought into interaction with each other in thought—*that* is what we call 'conceptual'.

We aim to understand reality, and such understanding is, in this rough-and-ready sense, conceptual. The philosophical discipline that is traditionally called *metaphysics* takes our search for an understanding of reality to its most general, encompassing scale; it asks fundamental questions that our ordinary and scientific investigations into reality often do not touch upon.[2] One such question of concern to the metaphysician, which is only implicit in our everyday and scientific efforts to understand this or that aspect of reality, is the following: what is the relation between our conceptual understanding of reality on the one hand, and reality itself on the other hand? This question motivates the present essay, which aims to be both a metaphysical investigation and an investigation into metaphysics. That is, it aims to provide answers to metaphysical questions, but also to contribute to a better understanding of metaphysics itself.

The issue of how our conceptual understanding of reality relates to reality itself is closely tied to central questions in other areas of philosophy. For instance, in *epistemology* one asks when a piece of putative conceptual understanding is justified, or under what conditions it qualifies as knowledge, and in the *philosophy of mind* one asks what it takes for there to be creatures that can thus conceptually understand anything in the first place. This essay deals primarily with the metaphysical question; we do touch upon these two related questions (and others), but only insofar as that is necessary or helpful for our metaphysical purposes.

Metaphysics, like all of philosophy, is a form of inquiry that focuses on thought, not observation. That does not mean, of course, that metaphysics is blind to the observational—it is a form of inquiry performed by *us*, beings endowed with faculties of both perception and thought, after all. Compare: we may one day reflect on our encounters with our neighbor and find ourselves considering quite general questions concerning human character, social interaction, friendship, and the like. Such reflection is itself conducted 'without observation'; we thereby engage in a non-empirical form of investigation, a purely 'conceptual' investigation in our intu-

[2] In the *Metaphysics*, Aristotle famously characterized metaphysics as 'the science of being *qua* being'—see [Aristotle 1998, pp. 79–80, 1003a21–31]. Stroud [2000, 2011] describes it as the 'Quest for Reality', Fine [2001, pp. 2 and 12] calls it the search for a '*metaphysical* conception of reality', a 'higher-order view of how the world really is', T. Nagel [1986] calls it the search for a 'view from nowhere' (but is critical about its prospects), and B. Williams [1978, p. 48–9] calls it the search for an understanding of 'absolute reality', reality 'as it is *anyway*'. Examples can be multiplied, of course.

itive sense. Metaphysics is non-empirical, purely conceptual, in the same way—but, likewise, it is of course triggered by our experiences.[3]

The attitude towards metaphysics as a philosophical discipline, throughout the centuries since its inception by Aristotle (and his precursors), has been quite diverse—ranging from enthusiastic speculation to extreme skepticism, from radically revisionary theorizing to stringent defenses of common sense, from consonance with to opposition to traditions of faith, and from consonance with to opposition to empirical science.

History also teaches that metaphysics is here to stay. Eventually, that most critical of episodes in the history of metaphysics, the skepticism mounted by the logical empiricists in the early twentieth century, was abandoned, giving rise to the return of metaphysics within the realms of their very heirs, the analytic philosophers. Even those defending the view that metaphysics should be 'naturalized' (that is, should be left to the empirical scientists) can agree that theirs is a methodological stance *about metaphysics*, not a rejection of the very discipline.[4] And similarly, those defending a strong anti-realist or skeptical position, according to which there is no true metaphysical theory to be had, assume that the search for an understanding of reality is intelligible—but, unfortunately, impossible.[5]

Metaphysics is, thus, alive and kicking—despite the fact, we should add, that more than 2500 years of philosophical struggle on the subject do not seem to have resulted in much progress, in the sense of an agreed-upon metaphysical body of categorical knowledge.[6] Reflection on this observation is likely to result in a pessimistic attitude towards attempts to make such progress. Be that as it may: our aim to understand reality still stands, and on the philosophical level it continues to find its expression in the contemporary metaphysical debate. Moreover, that debate is significantly different from what one finds in earlier centuries: for instance, analytic philosophy has developed its own approaches, marked by methods from logic and the philosophy of language, and scientific advances have put old metaphysical questions into a new light.[7] From this perspective, then, it makes sense to continue

[3] Frege remarks concerning what he takes to be the purely conceptual study of arithmetic and analysis: 'I do not mean in the least to deny that without sense impressions we should be as stupid as stones, and should know nothing either of numbers or of anything else' [Frege 1978/1884, §105].

[4] See, e.g., Ladyman *et al.* [2007] and Ladyman [2012].

[5] Some anti-realists disagree. We discuss their views in chapter 1, especially in §1.1.2.

[6] We write 'categorical' knowledge, because one could of course argue that a lot of *hypothetical* knowledge has been gathered—concerning what our world view should look like on the hypothesis that this or that metaphysical view is correct.

[7] For instance, the theories of relativity have challenged traditional metaphysical understandings of space and time, and evolutionary theory and molecular biology have challenged traditional metaphysical

approaching our quest for reality from a metaphysical perspective, starting from the contemporary discussion, and try our best.

A probing look at contemporary debates, however, may fairly quickly lead one to suspect that there are deep problems with the way in which their contributors proceed. The metaphysical discussion appears to be exceedingly fragmented: it is dominated by highly specialized contributions to many separate controversies, often without much feeling for the 'bigger picture'.[8] To someone not immersed in such controversies, it may well seem that the entire enterprise fails to connect to our quest for an understanding of reality—is it not merely an idle pastime, of dubious academic significance? Isn't what is going on mere speculation without a foothold in reality? Indeed, looking at the contemporary situation, one can easily get the feeling of being caught in a circle: renewed positive metaphysical efforts inevitably gives rise to renewed critical reflection on the very discipline. It is not surprising, then, that we have witnessed, in recent years, a return to the more skeptical attitudes towards metaphysics that prevailed in the beginning of the twentieth century—resulting in a familiar array of views. Those with a naturalistic bent of mind wonder: should questions concerning reality not be left to the experts, that is, to the scientists? And those with a more skeptical attitude ask: should we not finally accept that there is no true metaphysical story to be had? The recent rise of metaphysics within analytic philosophy thus brought with it a renewed reflection on the methodology of metaphysics: *meta*-metaphysics.[9]

As one comes to be familiar with the contemporary metametaphysical literature, one is bound to find oneself confronted with the very same fundamental question we formulated above, viewed now from a metametaphysical angle: *what is the relation between concepts and reality?* The realist metaphysician appears to think that his conceptual inquiries somehow have bearing on reality; the naturalist appears to think that such conceptual work is suspect unless it is firmly grounded in the empirical knowledge of the sciences; the anti-realist appears to think that there is an unbridgeable gap between (a putative) reality and the conceptual *as such*. As we will see, how one treats the question concerning the relation between concepts and reality determines to a considerable extent which metaphysical options one ends up with. The sketched three stances illustrate such determination of metaphysical views by one's take on the concepts/reality issue (as we will make clear in chapter 1).

understandings of living beings. We discuss these particular issues in §6.2.6 and in §7.4, respectively.

[8] Cases in point are the debates on persistence (see §6.2.1 and §6.2.5) and on material constitution (see fn. 33 on p. 24).

[9] See, e.g., the collection of essays published under the title *Metametaphysics* by Chalmers *et al.* [2009].

So much for an impressionistic sketch of the motivating question for our inquiry, and its background. Below, we briefly describe our aim, provide some remarks on the methodology we use, and then move on to give an outline of what is to come.

Aim and Method

The present essay deals both with metametaphysics and with metaphysics—it develops a metametaphysical approach that then is shown to give rise to a certain metaphysical outlook. The motivating question for the investigation is the one we formulated above: what is the relation between concepts and reality? This question has two quite different yet intimately connected readings: on the first, it asks whether and how metaphysics, a species of conceptual inquiry, can deliver anything having any bearing on reality at all. This is the metametaphysical or methodological reading. On the second reading, it asks what the metaphysical status is of the concepts by means of which we purportedly understand reality. This is the metaphysical reading.

The aim of our inquiry is to answer the motivating question on both of its readings. As it turns out, our answer to both readings can be stated by using the very same three-word phrase: *essences are concepts*.[10] This answers the first reading of our question: if successful, conceptual inquiry *is* metaphysical inquiry, for it is concepts that structure reality by being the essences of what is real. The same phrase also answers the second reading of our question: for everything that is, there is an answer to the Aristotelian question *what is it*?—and this answer is its essence, which is identical with the concept we aim to grasp when attempting to understand the thing in question.

Thus boldly stated, our thesis is bound to raise some eye-brows. What *are* essences? What does it mean to say that reality is 'conceptually structured'? The long answer to such questions is sketched in this essay as a whole; in short, however, we may perhaps clarify our statements with the help of the following comments. First of all, we should stress that, as we use the term, a 'concept' is not something merely in the head—if that is at all helpful, think of the medieval realists, who took universals to be 'in re' (as did Aristotle). Now take the concept *water*: that concept

[10] Although this introduction is not the right place for elaborate discussions, a few clarificatory remarks may be helpful here. We write 'essences are concepts': the thought is that all essences are concepts, or rather that essences are by nature concepts—not that all concepts are essences. There are very likely to be many concepts which are not essences, or at best only partly. Compare, e.g., the concepts *human being* and *bachelor*: the former is a good candidate for being an essence, the latter is not. Still, *bachelor* may be said to be *partly* an essence, since it arguably involves the concept *human being* as a part, in some sense of 'part'.

relates to a whole battery of further concepts, in specific, more or less interesting ways. Let us describe some of these conceptual connections: water is constituted out of hydrogen and oxygen in a 2:1 ratio; it can take solid, liquid, and gaseous forms; it can be used for culinary purposes; it plays an essential role for life on Earth; etc. Most of these conceptual connections can be enriched by adding explanations for them—that is, by clarifying why it is that the relevant connections hold. That is done by citing *further* conceptual connections. Suppose, for now, that we are correct in all of these conceptual connections—that is, assume that we have made no mistakes in grasping the concepts in question. It is crucial to notice that, by thus describing the relations between the concept *water* and other concepts, we are *ipso facto* describing *water*—the stuff itself. That is, on our supposition that we are not mistaken in our conception of water, water itself conforms to the described conceptual connections. This conformity on the part of something real (i.e., of water) to the relevant conceptual connections is what we have in mind when we say that reality is 'conceptually structured'. Our claim, now, is that this coincidence of the conceptual and the real holds because of the identity of concept with essence.[11]

Our essay is structured as follows. Part I introduces and defends a metametaphysical view, which we call the *Aristotelian picture*, that rests, centrally, on the thesis we have just introduced: that essences are concepts. Part I is divided up into three chapters: chapter 1 introduces the Aristotelian picture and sets it against the prevailing rival picture, which we call the Modern picture; chapter 2 sketches and defends the version of essentialism that it involves; and chapter 3 inquires into the relevant notion of conceptual truth. Part II then sketches the metaphysical view to which this Aristotelian picture gives rise, in large part by assessing key contemporary metaphysical questions and debates from its perspective. Part II is divided up into four chapters: chapter 4 introduces the basic conceptual framework of three conceptual 'gears' (see below) against the background of which we then discuss various topics from the contemporary metaphysical debate—Humeanism and laws of nature in chapter 5, causation and time in chapter 6, and life in chapter 7.

Our method is traditionally philosophical in nature, but slightly different for the two parts. Generally speaking, we are engaging in *a priori* conceptual investigations: we attempt to disclose important conceptual connections between basic

[11] Perhaps you think that the sketched conformity on the part of water is due to its underlying microphysical structure. That may well be so, but it would also illustrate rather than contradict our point: the concept of the relevant microphysical structure—H_2O, say—then supposedly contains what one needs to explain the mentioned conceptual connections. The conformity of stuff having just that microstructure to these (supposedly) deeper conceptual connections still illustrates our point.

metaphysical concepts, to discuss proposed construals of such concepts, to formulate our own proposals as clearly and unambiguously as possible, and to sketch the implications of such conceptual proposals—those of others and those of our own. In the first, metametaphysical part, we apply this general methodology in such a way as to uncover the metametaphysical underpinnings that are shared by the majority of contemporary metaphysical views (in particular, by metaphysical realists of various stripes and by metaphysical anti-realists). Insofar, our philosophical method is there applied in order to unearth the partly methodological assumptions of contemporary metaphysicians, identify problematic aspects of it, and propose a different methodology for metaphysics—a different framing of the metaphysical quest for reality. Consequently, the way we proceed in Part II partially depends on the results of Part I. In short, however, we may say already here that it involves identifying mutual conceptual dependencies amongst concepts belonging to certain identifiable clusters, explicating the metaphysical implications that the resulting observations have, and putting these metaphysical implications into perspective by examining them against the background of relevant contemporary metaphysical debates.

Having thus provided a sketch of what is to come in the most general terms, let us unpack these fairly dense formulations by walking through all seven of the chapters that this essay encompasses in an introductory fashion—first through those falling under Part I, then through those falling under Part II.

Outline of Part I: Metametaphysics

This part is devoted to introducing and defending the Aristotelian picture. The foundations of this metametaphysical picture are formulated in opposition to the widely accepted rival picture, the Modern picture, in chapter 1. Characteristic of this Modern picture is that it rests on a certain doubling of reality: the task of metaphysics, on this picture, is to provide an account of *reality as it is in itself* which should explain *reality as it is for us*.[12] The underlying thought is that reality-for-us is at least partly shaped by peculiarities of our own perceptual and cognitive make-up—in particular, by our concepts. This basic, metametaphysical orientation results in a duality of metaphysical options: metaphysical realism, which seeks to re-connect (part of) our conceptual grasp on reality with reality itself (for instance, through a correspondence theory of truth, or through truth-maker theory), and metaphysical anti-realism, which

[12]This doubling of reality is reminiscent of Sellars's famous distinction between the scientific and the manifest image—we briefly discuss its relation to the Modern picture in §1.2, p. 34.

takes this realist project to be futile, and instead resorts to a relocation of our aim to understand reality: we should be content with an understanding of a mere reality *for us*, and forget about the project of understanding reality as it is in itself. Metaphysics, on the Modern picture, is all about bridging (or not bridging) the very gap between concepts and reality that it rests upon.[13]

The Aristotelian picture, on the other hand, rejects the mentioned Modern picture distinction between reality-for-us and reality-itself, and thus rejects the separation of the conceptual from the real that this distinction invites as well. Or, positively put, the Aristotelian picture takes reality itself to be concept-involving, that is, it takes essences to be concepts, in a way to be spelled out. This spelling out is undertaken by reference to certain clusters of concepts in Part II of this essay, resulting in a sketch of the metaphysical view to which the Aristotelian picture gives rise. But first, we spell out and defend our claim that essences are concepts on a more general, still metametaphysical level, in chapters 2 and 3. The former is devoted to a defense of essentialism, the latter develops the crucial notion of conceptual truth. The notions of essential truth and of conceptual truth we thus arrive at turn out to coincide: that is how these two chapters underwrite the thesis that essences are concepts.

Our defense of essentialism, in chapter 2, starts with the observation that, quite often, the individuation of one or another kind of thing is taken to imply essential truths concerning instances of that kind. Thus, given that a set is individuated by its members (as the Axiom of Extensionality dictates), it is natural to conclude that it is impossible for a set to have different members from the ones it actually has: sets have their members essentially. We investigate the sources of this inferential move, which we call the *essentialist inference*, and conclude that it is not a purely formal-logical inference in the traditional sense, but rather rests on a certain understanding of *ontologically fundamental categories*, which we provide and defend. From a metametaphysical point of view, whatever one's metaphysical view looks like, it will include a list of fundamental sortal concepts (or at least a sketch of such a list), the ontologically fundamental categories, to which everything there is should be reducible—in a specific (and specified) sense of 'reduction'. The special ontological status of such categories grounds the applicability of the essentialist inference—and not only of that specific inference: it secures the essentiality of whatever follows from being an instance of the relevant ontologically fundamental category. It follows that essentialism is true whatever one's metaphysical views: whether essentialism holds

[13] F. Ellis [2005, p. 1] calls this way of separating the conceptual from the real, which she traces from the early modern period into the beginning of the twentieth century, 'the Syndrome'.

is not, in itself, an open metaphysical question.

The argument of chapter 2 is very general, and applies to metaphysical views both on the Modern picture and on the Aristotelian picture—yet it suits the Aristotelian picture best, and thereby lends it support. What might have seemed highly controversial when we first formulated it—that reality is essence-involving—turns out to be a default ingredient of any metaphysical view whatsoever. The complication that the Modern picture adds, which is its assumed task of bridging the very gap between concepts and (essence-involving) reality that it rests upon, is just that: an unnecessary complication.

Chapter 3, in turn, focuses on that key notion that occurs in our argument from chapter 2: what is essentially true of some thing are those truths that *follow from* the fundamental sortal concept under which it falls (the ontologically fundamental category). Such essential truths are, thus, *conceptual truths*. But what exactly does this 'following from' amount to? We said that the essentialist inference is not a formal inference in the traditional sense, hence the question arises how this 'following from', and the associated notion of conceptual truth, relate to the notion of formal-logical consequence as it is studied in the discipline of logic.

Logic can be understood as the study of consequence, of the very phenomenon of something following from something, which it studies through formalization in logical systems of various kinds. To that effect, it identifies the *logical form* of statements, and ignores their (further) content. Thereby, it restricts its attention to specifically logical versions of the consequence relation: those that can be crisply defined in the context of one or another logical system, based on one or another construal of logical form. Through such rigorous formalization, it is hoped, insight can be gained on valid inference in general—or, perhaps, it is hoped that scientifically workable versions of that intuitive notion can thus be identified.

We critically discuss a reductive approach to the relation of logical consequence in model-theoretic terms, pioneered by Alfred Tarski, and conclude that it fails in its reductive ambitions: when proposing a reductive definition of logical consequence, one is reasoning *about* that proposal in order to show that it is adequate—it 'follows from' the definition that the defined relation is truth-preserving, say. That reasoning itself relies on the notion that is allegedly analyzed (hence the scare quotes around 'follows from'). We observe, moreover, that what goes missing on such a reductive view is the distinctive modality of the consequence relation: whenever a conclusion follows from certain premises, the conclusion *must* be true if the premises are all true.

Having thus concluded that focusing on the formal approach to the consequence

relation does not illuminate the sort of following-from that we are interested in, we take a step back and consider the basic ingredients of that notion afresh: formality and modality. We argue that it is a mistake to think of logical form as *opposed to* content. Formality should not be understood in opposition to content, but rather as a restriction on content. The logical form of a statement is that which is left over if one considers only certain aspects of its overall content: the logical aspects (however one decides which these are). Thus understood, it becomes apparent that the fundamental notion is not that of logical form, but rather that of *truth guaranteed by content*—where 'content' now includes what was thought to be form. *Logical truth*, then, is truth guaranteed by the logical aspects of the content involved, whereas *conceptual truth* is truth guaranteed by all of the content involved. This puts the way in which logical studies contribute to our understanding of valid inference in general into a new light: such studies provide insight into the contributions that the specific, isolated aspects of content make to statements or thoughts. The other side of that logical coin is that an exclusive focus on those aspects of content that have been considered the 'logical' ones stands in the way of our recognizing *other* kinds of truth-guaranteeing content. Our proposal to take the notion of truth guaranteed by content as basic thus opens up the possibility of looking at the potentially truth-guaranteeing contributions that such complementary aspects of content, other concepts, make. And that is precisely what we propose to do in Part II, for certain metaphysically interesting clusters of concepts. Metaphysical truths are conceptual truths.

This, then, is the methodological basis for metaphysics that the Aristotelian picture embodies: we identify the kinds of contributions that the concepts we bring to bear in our understanding of reality make, and spell out the systematic metaphysical picture that we thus uncover—in particular, the variety of ways in which essences, which are concepts, work. It should be obvious, by now, that this construal of metaphysics is continuous with our every-day and scientific attempts to understand reality: it simply takes our enterprise of understanding reality to its most general conceptual level. It looks to the sciences and our every-day dealings with reality in order to find out which kinds of concepts our understanding of reality involves—in all of the aspects of it we care to notice.

Outline of Part II: Metaphysics

In this part, we move from the general, metametaphysical level to the metaphysical level, where we start to develop substantial metaphysical theses on the basis of the

Aristotelian picture as we sketched it in the first part. As our point of departure, we choose one inspiring thought which we borrow from Michael Thompson: that of a hierarchy of *conceptual gears*. The idea is, roughly, that there are certain clusters of concepts, focused around specific forms of predication, which are related in a peculiar way: when shifting from lower to higher conceptual gears, one shifts from simpler to more involved forms of predication, and with this increase in conceptual richness all of the concepts involved in undergo a certain kind of metamorphosis and differentiation too. Moreover, the concepts that belong to any one such cluster form a tight circle: none of them can be reductively defined in terms of lower-gear concepts.

By way of introduction, let us present the three conceptual gears that we discuss in the second part of this essay:

	First gear	**Second gear**	**Third gear**
form of predication:	tenseless	tensed/aspected	normative
applies to:	entity	object	organism
predicates involved:	feature	state/process	life-process
sortals involved:	type	natural kind	life-form
mode of being:	to exist	to persist	to live
engaged in:	—	process	life-process
characteristic:	geometry	causality/temporality	teleology/normativity
	Humeanism	*anti-Humeanism*	

This table merely gives a hint of what we intend to develop—but it illustrates the level of generality on which we are operating. The terms used are technical terms that require some explanation. The forms of predication form the back bone: on the first conceptual gear, we find a uniform, undifferentiated and in particular tenseless form of predication (roughly, Frege's notion of predication as function application); on the second conceptual gear we find a form of predication which is differentiated by tense and aspect, and on the third conceptual gear a further, normative differentiation is added. The individual things to which predication applies differ correspondingly among the conceptual gears: first, they occur as tenselessly existing entities, then as temporal, persisting physical objects, and then as living organisms. Similarly, the predicates that apply to such individual things differ among the gears. The ontologically fundamental categories (sortals) under which such individual things fall also transform when moving from gear to gear: we find types of entities, natural kinds, and life-forms. The category of activity is completely absent from the first conceptual gear, it appears on the scene when moving to the second conceptual gear

in the form of processes, and transforms in a peculiar way to become life-processes on the third conceptual gear.

In chapter 4, we introduce these three conceptual gears in much more detail, provide a sketch of the typical form of predication of each gear, and illustrate the typical forms of valid inference that each allows. We do not claim completeness; there may be further conceptual gears beyond or in between the ones we present, complete with their own forms of predication and forms of valid inference—but the exploration of these matters lies beyond the scope of our investigations.

The first conceptual gear applies to abstract objects, to the entities of mathematics and similar disciplines. The form of predication involved here is the simple, Fregean one that is devoid of complications such as tense, modality, and normativity. The attempt to use *only* this conceptual gear for understanding reality forms the basis for a specific version of Modern picture-based metaphysical realism: first-gear metaphysics. As such, first-gear metaphysics is committed to reductive views about all higher-gear phenomena. It is closely related to Humeanism, with its skeptical attitude towards reification of conceptual connections. We discuss this metaphysical stance in particular in chapter 5.

The second conceptual gear is marked by a form of predication that involves tense and aspect: time and causality are basic conceptual ingredients of this conceptual gear. Instead of mere entities, we have persisting physical objects, which not only bear features but are also engaged in activities—causal processes. The attempt to use *only* this conceptual gear (and the first one) for understanding reality forms the basis for another specific version of Modern picture-based metaphysical realism: second-gear metaphysics. As such, second-gear metaphysics is committed to reductive views about higher-gear phenomena; e.g., living organisms are to be understood, fundamentally, in mechanistic terms. Second-gear metaphysics is closely related to anti-Humeanism, with its opposition to Humean skepticism about 'necessary connections' and the like. We discuss this metaphysical stance, in its opposition to first-gear metaphysics, in chapter 6.

Finally, the third conceptual gear centers around a form of predication that involves not only tense and aspect but also a specific kind of normativity: the normativity involved in the teleology of life. Instead of mere physical objects, we are now dealing with living organisms, which not only bear features and can be involved in second-gear processes, but are also engaged in third-gear life-processes, which have a natural end—the well-being of the organism in question. The last chapter of this essay, chapter 7, discusses this conceptual gear, and in particular defends the au-

tonomy and irreducibility of this conceptual gear against the prevailing second-gear metaphysical consensus.

Our discussions in chapters 5, 6, and 7 of Part II pursue two aims at once. On the one hand, they aim to spell out, in more detail than the global sketch of the introductory chapter 4 allows, what exactly the import of the conceptual gears is. This is especially the case in chapters 6 and 7. In the former, we argue that the second conceptual gear includes a commitment to a powers-based understanding of causality, and to a presentist, A-theoretic and endurantist understanding of time. In the latter, we argue that third-gear phenomena—that is to say, life—presupposes an open future in the second-gear sense, and entails a fairly strong form of essentialism concerning life-forms (whose relation to biological species is, however, not entirely straightforward).

On the other hand, however, our aim in these three chapters is to elucidate how Modern picture-based metaphysical realism, in the guise of first-gear metaphysics and second-gear metaphysics, tends to distort the metaphysical project. For instance, in chapter 5 we illustrate, by reference to the metaphysical program of David Lewis, the first-gear metaphysicians' habit of providing reductive analyses of higher-gear phenomena: by doing so, it seems that they can have all the higher-gear cake they want and eat it, yet on closer inspection their cake is baked from bland, first-gear conceptual ingredients only—they merely pay lip service to the higher-gear aspects of reality. And, more particularly, we survey key contemporary metaphysical debates from the perspective we developed, especially insofar as they portray clashes between Humeanism and anti-Humeanism (that is, between first- and second-gear metaphysics). In §5.2, we show that attempts to provide a satisfactory non-governing, Humean understanding of laws of nature is indeed a first-gear enterprise, and as such is committed to the entire reductive program that it embodies. In §6.1, we discuss attempts to provide an anti-Humean understanding of causality (a second-gear enterprise), and conclude, for instance, that the 'necessitarian' view defended by David Armstrong collapses into a first-gear surrogate of causation because his overall metaphysics is clearly first-gear. The only way of endorsing the second-gear notion of causality is by adopting the entire second conceptual gear, which includes the concept of a causal power. And in §6.2, by far the longest investigation within the present essay, we discuss the search for a non-first-gear understanding of time: it illustrates how intricate first-gear metaphysics can get if it is combined with a serious attempt to capture what the second-gear thinker requires. In particular, it illustrates that even those who aim at capturing a truly second-gear understanding of time in

fact may end up with a particularly subtle but still first-gear substitute—just like what we concluded with regard to Armstrong's necessitarian view. Moving beyond the Humeanism/anti-Humeanism dispute, large parts of chapter 7 are devoted to an attempt to make clear that the motivation for the reductive stance on the realm of the living, which appears to be widespread even amongst those defending one or another version of anti-reductionism in biology, crucially rests on endorsing a Modern picture-based metaphysical realism of the second-gear variety. Although much less visible as such, the issue of reductionism in biology is structurally very similar to what is at stake between the Humeans and the anti-Humeans, but one step higher on the ladder of gears.[14]

The picture that emerges throughout our discussions is that of a reality that is at least as complex, multifaceted, and rich as the conceptual resources we employ in our quest to understand it. It is not the abstract, almost mathematical realm of which first-gear metaphysics dreams: it also includes concrete, persisting physical objects with their powers and the causal processes they are engaged in. Moreover, it is not merely the purely mechanical world that second-gear metaphysicians take it to be either: it also houses living organisms that organize the stuff they are made of in accordance with their nature by way of the teleological life-processes they are engaged in. And, although this lies beyond the scope of the present inquiry, one may envisage that the list of conceptual gears does not stop here: on reflection, sentient and rational life may well deserve further conceptual gears of their own. On the Aristotelian picture, then, reality turns out to be a place in which we may find ourselves to be at home.

* * *

[14] Notice that a Modern picture-based metaphysical realist might agree with us on the irreducibility of the third conceptual gear—he might in fact agree with the bulk of Aristotelian picture-based metaphysics, despite holding on to the Modern picture. In such a case, however, the Modern picture distinction between reality-for-us and reality-itself collapses—and thereby the whole point of endorsing the Modern picture.

Part I

Conceptual Realism: Metametaphysics

Chapter 1

The Aristotelian Picture*

In this chapter, we introduce the *Aristotelian picture*. It is Aristotelian not in any exegetical sense, but rather Aristotelian in spirit, as will become clear along the way. Our very first task is to sketch what is meant by 'picture', and to clarify what the role the Aristotelian picture plays for this essay is.

The Aristotelian picture is a metametaphysical view centered around the basic insight that essences are concepts. We will slowly develop what we have in mind below. Our strategy is to set the Aristotelian picture against the prevailing metametaphysical picture which we call the *Modern picture*. We extract our sketch of that Modern picture, in turn, from an investigation into contemporary metaphysical realism and metaphysical anti-realism. This procedure also brings to light important reasons for abandoning the latter picture in virtue of the former, although these reasons will not even begin to be conclusive until a more rigorous exposition and defense of the Aristotelian picture has been given. The remainder of Part I of this essay—chapters 2 and 3—aims to contribute to this exposition and defense on a metametaphysical level, whereas in Part II we investigate its metaphysical implications.

Our enterprise connects to various attempts one finds, in recent years, to overcome difficulties of the Modern picture by questioning that very picture—though, of course, not in these precise terms. Several philosophers, such as John McDowell and Hilary Putnam (but also, though less explicitly, David Wiggins, Fiona Ellis, Sebastian Rödl and Michael Thompson) have, from various philosophical angles, contributed to this trend.[15] The present essay in general, and this first chapter in particular, aims to make

*Large portions of this chapter have been drawn from Mulder [2012].

[15]See, e.g., McDowell [1994a, 2009a,b], Putnam [1992, 1999], Wiggins [2001], F. Ellis [2005], Rödl [2007,

progress on this score by exploring an alternative to the Modern picture specifically from the point of view of (contemporary analytic) metaphysics.

In thus providing a first sketch of the place of the current chapter within the overall structure of this essay, we speak of 'metaphysical' and of 'metametaphysical' investigation. As with ethics and meta-ethics, the distinction is one between a certain type of investigation on the one hand, and reflection on that type of investigation on the other; between attempting to answer metaphysical (or ethical) questions and considering whether and, if yes, how such questions can or must be answered. And, as with meta-ethics, a fundamental divide can be discerned on the metametaphysical level between realists and (global) anti-realists.

Within analytic philosophy, an increasing interest in metametaphysical questions can be observed—which was only to be expected, given the remarkable resurrection of metaphysics in recent decades.[16] Within the metametaphysical discussion, some complain that one or another metaphysical dispute is 'merely verbal'[17], others develop theoretical frameworks designed to show that one or another metaphysical view fares best with respect to some allegedly universal standard for metaphysical theorizing (familiar theoretical virtues such as parsimony, simplicity, strength, and fit are frequently alluded to)[18], still others stress the importance of our scientific and/or common sense understanding of reality as a guide for proper metaphysical theorizing.[19] Looking at what is going on both on the metaphysical and on its meta-level, we find the mentioned familiar and fundamental division between what we call *metaphysical realists* and *metaphysical anti-realists*. Realists understand the metaphysical project to be about capturing the true, objective, mind-independent structure (or make-up, or layout, or foundation) of reality—in short, realists aim at a theory that captures reality as it is *anyway*, to use Bernard Williams's apt phrase.[20] Anti-realists, on the other hand, insist that this project is problematic, or even incoherent, roughly because we always approach reality from a certain perspective, a certain conceptual

2012/2005], and M. Thompson [2008]. Looking beyond contemporary philosophers, sources of inspiration for our opposition to the Modern picture have been Hegel [1977/1807, 1969/1816], Steiner [1886, 1894], and Husserl [1970/1954], and, of course, Aristotle [1963, 1984, 1996, 1998].

[16] See, for instance, the recent collection of essays entitled *Metametaphysics* [Chalmers *et al.* 2009].

[17] See especially Hirsch [2002a,b, 2005, 2008], though his complaint is not aimed at metaphysical disputes generally but rather at quite specific ones.

[18] See, e.g., Rodriguez-Pereyra [2002, ch. 12] and Sider [2012, ch. 1], and also Sider [2009] and other essays in the mentioned *Metametaphysics* collection [Chalmers *et al.* 2009]. See also §1.1.1 below.

[19] For instance, see, respectively, Ladyman *et al.* [2007], Callender [2011a], and Ladyman [2012]; and Elder [2005, 2011].

[20] See B. Williams [1978, pp. 48–9].

point of view, which makes an understanding of reality at all possible and hence cannot be 'peeled off' in order to reach this alleged objective, mind-independent reality. Hence anti-realists suggest that we should give up on this quest entirely.[21]

This brief sketch of the realism/anti-realism divide makes salient a third option: a view which takes the project of the realist seriously, but finds the anti-realist's arguments against its viability compelling. One may describe this view as metaphysical pessimism—and in §1.4 below we will see that it is, in fact, a type of skepticism. Such metaphysical pessimism will only start to play a role later in this chapter; we start out by considering metaphysical realism and anti-realism.

Among the many confusions surrounding the 'realism'/'anti-realism' terminology there is one we should remove straight away. Whenever we talk about anti-realism, we mean global, metaphysical anti-realism, as opposed to local anti-realism. To illustrate, consider the ethical anti-realist: he thinks that, in reality as it is *anyway*, there is no normativity of the kind the ethical realist endorses; instead, that kind of normativity somehow arises out of our attitudes towards actions and practices (say). Such local anti-realisms only make sense against a background of metaphysical realism: in reality, there are only natural, non-normative matters of fact (the realist metaphysical background), hence normativity is to be accounted for by way of the ethical anti-realist's preferred story (the local anti-realism).

Local anti-realism is, thus, metaphysical realism. Global, metaphysical anti-realism, on the other hand, rejects the very business of establishing what reality is like in itself in order then to exclude specific ingredients from it. It gives up on this project entirely. Continuing our example: for the metaphysical anti-realist the normative and the non-normative alike are the result of our conceptualization of reality, in normative and non-normative terms, respectively. Acceptance or rejection of either kind of conceptualization is to be argued for on pragmatic grounds, not on metaphysical grounds.

Although metaphysical realists and anti-realists *seem* to embrace almost entirely opposed metametaphysical as well as metaphysical views, they turn out to share a philosophical core, a metametaphysical picture, which is presupposed by the very dichotomy of realism and anti-realism. In this chapter, we extract from key observations within the contemporary metaphysical and metametaphysical discussions

[21] It is difficult to characterize the realism/anti-realism divide in a neutral way, without already offending one or another defender of a version of (anti-)realism. The purpose of our present characterization, however, is merely to sketch the divide, not to offer a precise formulation. In §1.1 below we develop a more detailed exposition of the divide.

precisely that underlying picture: the *Modern picture*.[22]

Along the way, we point out certain problematic features of both metaphysical realism and the metaphysical anti-realism that turn out to be symptomatic of the Modern picture. When taken seriously, these problematic features tend to result in a form of metaphysical pessimism: we aim to arrive at a certain understanding of reality, yet we find that progress on that quest is blocked by grave, perhaps even insurmountable obstacles. On reflection, these observations motivate the search for an alternative approach to metaphysics, one which does not rest on the Modern picture. This may seem to be a fairly traditional quest: to find a 'third way' that is neither metaphysically realist nor metaphysically anti-realist in spirit. In a way, that is precisely what we will be gesturing towards in this chapter, but we should keep in mind that this 'third way' is not really a third or middle way between metaphysical realism and anti-realism, since it rejects the very picture on which the latter two depend. Put very briefly, our alternative picture, the *Aristotelian picture*, centers around the insight that essences are concepts, whereas the Modern picture postulates, in effect, a strict distinction between the conceptual and the real. That is why we have provided our essay with the title *conceptual realism*—but we should not get ahead of things too much.

We proceed as follows. The first section (§1.1) is devoted to a critical discussion of both the realist and the anti-realist positions and their metametaphysical underpinnings, leading to a sketch of the mentioned Modern picture.[23] Our discussion of realism focuses on one particularly telling feature: it yields a growing range of radically diverging views between which it is impossible to adjudicate on the basis of the realist's metametaphysical understanding of how their project is to be carried out. There is just no way to put aside even the most outrageous of metaphysical speculations—a good testimony for which is the continuing recurrence of such positions (e.g., modal realism, blobjectivism, nihilism[24]). Our discussion of anti-realism in turn focuses on one of its most striking features: it aims to reject the very idea of a mind-independent reality, yet depends crucially on that same idea for its own articulation. We focus more on realism than on anti-realism, partly because it takes

[22] We adopt the metaphor of a 'picture' from the later Wittgenstein, who famously remarked regarding his own earlier (*Tractatus*) views that 'a *picture* held us captive'—see Wittgenstein [1953, §115] (and [1921]). However, we do not mean to suggest affinity with his quietist, anti-metaphysical project of bringing 'words back from their metaphysical to their everyday use' [§115]—see also fn. 72 on p. 52 below.

[23] For simplicity's sake, we will often drop the qualification 'metaphysical' and just talk about realism and anti-realism.

[24] See Lewis [1986a] and Cameron [2012] for modal realism; Horgan and Potrč [2008] for blobjectivism; Unger [1979a,b,c, 1980], Dorr and Rosen [2002], and Dorr [2005] for nihilism. See also §1.1.1 below.

a bit more effort to unearth the metametaphysical background of the realist project, and partly because we will have occasion to revisit several realist metaphysical views in the second part of this essay.

The second section (§1.2) develops our diagnosis of the sketched situation, which comes down to spelling out the defining features of the Modern picture. The very basis of the Modern picture turns out to be its strict separation between reality as it is for us and reality as it is in itself. We trace out in some detail how this picture gives rise to the unsatisfactory duality of (meta)metaphysical options problematized in §1.1. Along the way, we come to see that against the background of this Modern picture, the problematic aspects of realism and anti-realism highlighted in the previous section are to some extent inevitable. We mention that there is a way out, for the adherent of the Modern picture: it consists in endorsing a form of skepticism. On such a view, the realist project is retained, but its problematic aspects, along with the anti-realist's objections to it, motivate a pessimistic estimate as to the prospects of metaphysical progress.

Before discussing such metaphysical pessimism, we introduce, in section §1.3, by reference to the defining features of the Modern picture as developed in §1.2, our alternative metametaphysical picture: the Aristotelian picture. We start by rejecting the strict separation between reality as it is for us and reality as it is in itself which defines the Modern picture, and trace out the consequences of this thought, which leads us to put forth our central thesis, to be developed further in chapters 2 and 3: that essences are concepts.

Having thus sketched the metametaphysical picture which this essay is devoted to, we move on to discuss, in §1.4, the metaphysical pessimism mentioned above. To that effect, we critically examine Barry Stroud's conception of metaphysics: we discuss his pessimistic estimation as to the prospects of metaphysical inquiry and argue that this is indeed due to his being implicitly committed to the Modern picture, despite being critical of both metaphysical realism and metaphysical anti-realism. Thereby, we further illustrate both our reasons for abandoning the Modern picture and our conception of the alternative Aristotelian picture. A brief summary concludes this chapter in §1.5.

1.1 Uncovering the Modern Picture

Realists attempt to construct metaphysical theories purporting to describe the fundamental/objective/mind-independent structure or make-up of reality (or something

such).[25] Anti-realists attempt to show that realists embark upon a futile quest, that the only sensible notion of 'reality' we can have is that of a reality as carved up by us, that is, a reality that is partly constituted by our own conceptual activity, our cognitive perspective (or something such). One popular way of expressing the pivotal point of disagreement between realists and anti-realists is as follows: realists urge that their theories describe the world as it *mind-independently* is, whereas anti-realists argue that the very categories employed in such purported descriptions are, after all, *our own* categories, and hence that the very idea of a mind-independent reality in the realist's sense is deeply mistaken.

Before we start, a cautionary remark is in order. In this section, we move fairly quickly through the complex and intricate fields of metaphysical realism and anti-realism. Our objective is not to achieve completeness or conclusiveness, but rather to arrive at a reasonably clear general framework that allows us to discern the shared metametaphysical orientation of both camps: the Modern picture. We are thus going to make a few quite sweeping claims about realism and anti-realism, without developing the thorough support such claims require in much detail. Such more rigorous and detailed assessments of a variety of metaphysical views will be presented in Part II of this essay, but those require that we first develop a general sketch of the Modern picture and our alternative, Aristotelian picture. That is what we aim at in this first part.

1.1.1 The Metametaphysics of Realism

Realists hold widely diverging views. We mention four of the more well-known positions on the metaphysical battle-field. There is the camp of David Lewis, whose broadly Humean metaphysical program includes four-dimensionalism, Humean supervenience, and counterpart theory.[26] There is the camp of David Armstrong, whose metaphysical views center on universals between which necessary connections hold and which are instantiated by objects, thus resulting in states of affairs, which serve as truth-makers for our claims about the world. Armstrong, like Lewis, has defended his metaphysical views with exceptional vigor and clarity over a considerable number

[25] Jenkins [2010] contains a helpful discussion of various ways of spelling out realism. She recommends construing realism in terms of objectivity or mind-independence, as we do, and distinguishes realism in this sense both from the claim that ontological disputes are substantial ('inflationism') and from the claim that there is an ontologically weighty reading of the quantifiers ('quantifier invariantism').

[26] See, e.g., Lewis [1986a, 1991] and Sider [2001, 2012]. This Humean metaphysical program will occupy us briefly in §2.3 and, at greater length, also in chapter 5.

of decades.[27] There is the currently quite popular camp of the dispositionalists, who object to the overly static, 'powerless' conception of properties of both the Lewisians and the Armstrongians, and defend a metaphysical view centered around the notion of a power or disposition.[28] And then there are those who regard all of the above with suspicion or even strong disapproval because it is the sciences, in their view, that decide what reality looks like—still, they go on to defend a metaphysical view based on their understanding of what science (or, more specifically, physics) tells us, resulting in views like structural realism (here we wander off from contemporary metaphysics into the philosophy of science—although the debate about scientific realism is itself, of course, a metaphysical debate).[29]

The list could be considerably extended, especially if one were to include not only the more comprehensive and mainstream views but also the more local and exotic ones. And it is of course an illusion to think that the listed families of views are cleanly separated: one can find metaphysicians defending almost any combination of aspects of any view on the list.

A brief note aside is in order. Perhaps the most salient omission in our above list of metaphysical camps is the camp of the neo-Aristotelians—are these not clearly to be included under the heading of metaphysical realism? The problem is that many neo-Aristotelian views turn out to fit poorly into the metaphysical realist's camp, because they are in tension with the underlying Modern picture, although not strictly incompatible with it. It is no coincidence that we label our alternative picture the Aristotelian picture.[30]

Now, given this diversity of views, and the fundamental level at which disagreement starts, it is not surprising that many discussions within the realist camp end in impasses. This is an interesting and characteristic feature of realist metaphysics, and a phenomenon worth studying with some care. Of course, divergence in views is the usual situation not just in realist metaphysics but in philosophy more generally—but, we will claim, the divergence one can observe within realist metaphysics is special

[27] See Armstrong [1978a,b, 1983, 1997, 2004b, 2010]. We discuss Armstrong's metaphysical position at some length in chapter 6 (see esp. §6.1 on causation).

[28] See, e.g., Heil [2003], Mumford [2004], Mumford and Anjum [2011], Molnar [2003], and Bird [2007], and the essays in Marmodoro [2010] and Greco and Groff [2013]. We will have a closer look at such metaphysical views in our discussion of causation in §6.1.

[29] See especially Ladyman [2002], Ladyman *et al.* [2007], and the essays in Landry and Rickles [2012].

[30] See, e.g., Wiggins [1995, 2001], Lowe [2006, 2008b, 2009], Fine [1994b, 2003], Oderberg [2007], and the essays in Tahko [2012]. For some of these philosophers, one may perhaps show that they are already endorsing the Aristotelian picture rather than the Modern picture. But not for all of them—see also fn. 64 on p. 44 below.

in that it betrays an underlying metametaphysical tension.[31]

Let us see whether we can understand this divergence phenomenon on the basis of a metametaphysical reflection on the very project of the realist metaphysician. Realist metaphysics is normally conceived of as, at root, an *a priori* discipline, and hence the only way to adjudicate between the diverse positions is by examining various features of the metaphysical theories themselves.[32] Typically, simplicity, parsimony, strength, coherence, and fit are amongst the features that are deemed of importance: a metaphysical theory should not be needlessly complicated; it should not postulate entities without sufficient reason for doing so; it should be comprehensive, accounting for many aspects of reality at once; it should be both internally coherent and consonant with relevant established further theories; and it should be consonant with the relevant aspects of our scientific and/or commonsense understanding of reality.

Such theoretical virtues thus constrain metaphysical theorizing in an *a priori* manner, by favoring theories that score better on the virtues over others, but also in an *a posteriori* manner, by favoring theories that do not entail too much of a revision with regard to our scientific and commonsense understanding of reality. It is assumed, then, that the world as it is according to our scientific/commonsense understanding of it is in need of a specifically metaphysical explanation (even if it is then urged that this explanation should also come from the sciences). Typical test cases for metaphysical theories include the traditional metaphysical puzzles of identity, composition, persistence, etc.[33]

The overall strategy is, thus, like that of an inference to the best explanation: whichever theory explains best how the world is *for us*, as we know it through science and/or everyday life, whilst using as little and elegant (etc.) resources as possible wins the metaphysical prize of capturing how the world is *in itself*.[34]

[31] In the literature, one finds that such impasses are often noticed by anti-realists, and used in their arguments against metaphysical realism. See §1.1.2 below, and references there.

[32] See the references provided in fn. 18 on p. 18 above.

[33] There is a lively literature on such puzzles and the ways in which different metaphysical views deal with them. See, e.g., Geach [1980], Wiggins [1968, 2001], Gibbard [1975], Thomson [1983, 1998], Burke [1994], Fine [1994a, 1999], Rea [1995], Sidelle [1998], Sider [2001], Merricks [2001], Koslicki [2004, 2008], and Paul [2006, 2010].

[34] Notice that what is to be explained by such a metaphysical theory is not the way the world *appears* to us but rather the way the world *is* for us (which includes mere appearances as well as concrete physical things). As said, scientifically-minded realists tend to import scientific theories into their metaphysics; some of them go so far as to pronounce metaphysics 'naturalized' [see Ladyman *et al.* 2007]. Here is a recent example: Williamson [2013] praises, in the context of his peculiar defense of 'necessitism' (the view that, necessarily, everything necessarily exists), 'the advantages of an abductive methodology . . . in clarifying the relative strenghts and weaknesses of contingentism and necessitism' [p. 288], and urges that these rival views in modal metaphysics 'need to be compared on the usual dimensions for theory choice

There are quite some problems with such an approach based on theoretical virtues, as is well known: disagreements amongst philosophers usually involve disagreements as to which theory scores higher on one or another theoretical virtue. This is only to be expected—for instance, there is lots of room to shift the priorities: is a parsimonious but more complicated theory better than a simple but less parsimonious one? And: why are *these* the important theoretical virtues, and not others? We will not discuss such questions here. Rather, we aim to precisify our observation of divergence within metaphysical realism by arguing that even if such problems of choice of theoretical virtues and their comparative weights can indeed be overcome, as many realists think[35], there is still no reason to think that the noticed divergence will be any less radical.

Consider the following theory, which we call *numberism*: the only things that exist are the real numbers, with their many properties and the many relations that hold between them. We construe these properties and relations as sets of (n-tuples of) real numbers. That's all: a very simple and parsimonious theory, and strong too, since there is no truth about reality that is not implied by it. You can say very many true and false things about the real numbers, so the only thing we need is a suitable map from ordinary and scientific truths (and falsehoods) to those truths and falsehoods. Such a map, which guarantees that our theory of numberism also scores high on the theoretical virtue of fit, can be arranged for by way of 'indirect correspondence' (a strategy actually used by Horgan and Potrč [2008, see esp. ch. 3]): the truth of scientific and/or ordinary truths consists in no more (and no less) than that certain complicated statements about the real numbers are true. For example, 'you are now reading this essay' would be true, on such a view, just if there are three real numbers which stand in a specific ternary relation. It surely is no easy task to provide a *systematic* way of paraphrasing every true sentence into some sentence concerning the real numbers, but there seems to be no reason why it could not be done (e.g., with the help of a Davidsonian truth theory).[36]

Numberism is false. And not only that, it is an outrageous metaphysical theory. Yet the metametaphysical orientation of the realist debate in metaphysics, as we have just provisionally sketched it, provides no real grounds for dismissing it. Indeed, this orientation is more likely to favor numberism over many alternatives that are actually defended, given its excellent score on simplicity, parsimony, and strength.

in science'—he mentions strength, simplicity, symmetry, and elegance [p. 276–7].

[35] See, e.g., Rodriguez-Pereyra [2002, ch. 12] and Sider [2012, ch. 1].

[36] Quine actually considered a theory very much like numberism once; he rejected it mainly because the ideology would be impractical [Quine 1964].

Numberism explains what needs to be explained: the way the world is for us; and it does so by using only very simple, parsimonious and strong resources. The theory may be counterintuitive, but that does not prevent it from being just as worthy a candidate as any other metaphysical theory.

Realists are likely to object to our considerations so far by pointing out that we have been cheating when it comes to the theoretical virtue of fit: using 'indirect correspondence' to secure consonance with our ordinary and/or scientific convictions in the way we do fails entirely to do justice to the point of this virtue. Well, here we can agree: if consonance with scientific and/or commonsense wisdom can be gained in such cheap a way, there is no reason for establishing it at all (except for rhetorical reasons). The question then becomes: what *more* is required of a metaphysical theory in order for it to score high on the virtue of fit? The answer is likely to be something like this: we should take our scientific and/or commonsense knowledge *seriously*.

If, after due examination, we declare that there is an oak tree in the backyard, this means that there really is an *oak tree* in the backyard (barring skeptical scenarios), and so oak trees should be allowed to be part of the furniture of the world—thus ruling out numberism, which does not include oak trees in its ontology, but only real numbers. But then, what *are* oak trees, metaphysically speaking?

Here we get to the heart of the matter: in reply to such a question, the realist allows himself to say *anything* whatsoever—even that oak trees really *are* numbers. To switch to a more serious example, take the following Lewis-like view. Oak trees are space-time worms (perdurantism), that is, mereological sums of space-time points displaying certain patterns of qualities. Now, let us say that space-time points may be thought of as quadruples of real numbers (coordinates), and the properties instantiated at those points may be thought of as sets of such quadruples: then, oak trees can be thought of as sets of quadruples of real numbers that are members of certain properties. That is pretty close to numberism, even though we started out by insisting that we take our scientific and common sense knowledge 'seriously'. Of course, this argument is not intended to show that Lewisian metaphysics just is numberism, because it is evidently not.[37] The point is, rather, that if such a Lewis-style conception of physical objects is acceptable from a realist metaphysician's point of view, then numberism is too.

It seems, then, that the metaphysical realist is in no position to make our commonsense and/or scientific understanding of reality have any constraining effect on

[37] Although there really is a Lewisian metaphysician who comes very close to defending the sketched Lewisian view—see Heller [1990]. We discuss Lewisian metaphysics briefly in §2.3, and at greater length in §5.1.

metaphysical theorizing. As a result, wildly diverging metaphysical views are being developed and defended, some bordering the level of outrageousness we attempted to illustrate with our example of numberism (we have mentioned a few of them in passing). Hence it becomes impossible to decide which of the many alternative metaphysical theories is the correct one—impossible even to dismiss such outrageous theories as numberism. This, in turn, leads to the observed impasses within the metaphysical debate, and explains the dissatisfaction that many philosophers have expressed with regard to analytic metaphysics.[38]

Of course, the considerations we have voiced at best hint at a case for these bold conclusions—a proper argument would have to dig much deeper into the various versions of realism. That is not our aim here (though Part II contains some relevant investigations); our survey of realism (and anti-realism) only serves to expose the underlying metametaphysical picture. We can now start to formulate that picture: as realists, we take ourselves to be discussing the ultimate make-up of reality, but we are supposing this ultimate make-up to be something that lies *behind* the world as it is known to us through science and/or common sense—indeed, it is supposed to explain this. Because of this set-up, there simply *is* no decent way of doing justice to the point of the theoretical virtue of fit: what the metaphysical theory is supposed to capture is from the very start conceived to be entirely disjoint from reality as it is for us. The picture that is assumed is that familiar ever since it became popular during the Early Modern period: there is this 'something I know not what' that lies beyond the world as it is *for us*, and this something is the target of the realist.[39] Given such a picture, it is not at all surprising that there is so little convergence amongst realist views. We are locating the target of those views in an unreachable place, as it were. One can see why Kant decided that there really is no point in speculating about what such a 'world-in-itself' is like—but we should not get ahead of ourselves.

We may illustrate the point we have been developing by looking at Fraser MacBride's review of Rodriguez-Pereyra [2002], who defends a version of 'resemblance nominalism'. MacBride opens that review with a very striking statement:

> Take David Lewis' theory that possible worlds are disconnected spatio-temporal regions whose inhabitants we routinely discuss in ordinary modal discourse. Fa-

[38] See, e.g., Ladyman *et al.* [2007], Callender [2011a], Ladyman [2012], Maclaurin and Dyke [2012], and Dyke and Maclaurin [2013].

[39] For the phrase 'something I know not what' see, e.g., Locke [1790, Book II, ch. xxiii, §15]. Philosophers like John Locke were very fond of the corpuscularian speculations of the physics of their times, of course, and it is similarly open to the contemporary realist to hold that present-day physics has a metaphysical say on what reality-itself looks like. We already remarked that this thought motivates certain realist philosophers of science, such as Ladyman *et al.* [2007] and Ladyman [2012].

> mously, Lewis' theory met with an incredulous stare from those to whom it was presented. But Lewis faced the incredulous stare down. Following arguments where they led, he showed that his theory enjoyed benefits that outweighed the cost of offending intuition. In doing so Lewis exercised a liberating effect on contemporary metaphysics, blowing away the cobwebs of ordinary-language philosophy and making it acceptable for philosophers to advance bold and surprising claims about reality. [MacBride 2004]

What we have been working towards is this. It is fine that Lewis has succeeded in liberating us from 'the cobwebs of ordinary-language philosophy', in getting us to critically reflect on deeply-ingrained prejudices and intuitions concerning metaphysical questions—but, by 'making it acceptable to advance bold and surprising claims about reality' he also contributed to opening the floodgates to a never-ending stream of ever-more-diverging metaphysical views. That is a symptomatic problem of contemporary metaphysical realism.

To repeat: these brief and sketchy remarks are not meant to provide a conclusive *argument* against metaphysical realism. We are highlighting features that are symptomatic for the metametaphysical picture on which the project of metaphysical realism is based. That picture, which we call the Modern picture, delineates the target of the realist project in such a way that the disconcerting divergence we have highlighted results. The target is located *behind* the reality we are familiar with, so to speak, and hence there is nothing to constrain how one goes about conceiving of this metaphysical underpinning of reality. The lack of such a constraining standard is a flaw in the design of the very project of metaphysical realism. To repeat: our aim in presenting these considerations is merely to motivate the critical exploration of the underlying metametaphysical picture, the Modern picture, in order then to move on to the alternative picture to which this essay is devoted: the Aristotelian picture. But first, we now turn to anti-realism.

1.1.2 The Metametaphysics of Anti-Realism

The anti-realist is impressed by the fact that whatever theory we come up with, it will be couched in a certain language, it will depend on a certain conceptual scheme. This insight shapes his understanding of metaphysics, or rather, his rejection of metaphysics as the realist conceives of it. Hence, he will not be surprised that there is so much disagreement amongst realists: endorsing different metaphysical theories is, from an anti-realist point of view, much like viewing reality from different conceptual points of view. Hence, nothing *can* constrain the realist's metaphysical theorizing. Yet for the anti-realist this doesn't mean that metaphysics is useless—as Carnap

famously argued in his [1950], we should simply locate the significance of such theory construction elsewhere. Instead of asking which theory is true, we should ask the pragmatic question: which theories are *useful*? Does it help us to adopt one or another scheme of thinking? From this point of view, there is no absolute reason to change our language or our conceptual apparatus, nor is there any reason to choose one theory to the exclusion of all others: pluralism makes more sense, since different theories might well be useful for different purposes.[40]

From such a perspective it seems indeed as if all those quarrels among realists are merely 'verbal disputes'[41]: it's all just a matter of 'language choice', or choice of conceptual framework. For anti-realists, doing metaphysics is nothing more than discussing which conceptual schemes to use for what purposes. For example, if we find possible-world talk useful, we adopt it. The question as to whether possible worlds really exist or are rather fictions or something such simply does not arise. Those are, to use Carnapian terminology, *external questions*, as opposed to *internal questions* phrased within the conceptual scheme we have adopted (e.g.: what should the accessibility relation look like when explicating alethic modality in terms of possible worlds?).[42] Since many of the theories that realist metaphysicians have to offer are precisely aimed at such external questions, it will not be surprising that these are simply not interesting from an anti-realist point of view.

Anti-realism thus takes a deflationary stance towards metaphysics. But, we may ask, at what cost? What exactly does anti-realism have to presuppose in order arrive at such a deflationary stance? Once we attempt to spell out these presuppositions, we stumble upon some uncomfortable aspects of the anti-realist family of views, three of which are useful for our purposes of uncovering the metametaphysical picture on which anti-realism rests.

Firstly, and most importantly, observe that the anti-realist wants us to simply forget about reality-in-itself, which is the target of realist metaphysics, but still requires this idea in order to state his very position: we should be concerned with reality as it is for us *as opposed to* reality as it is in itself. One topic for which this becomes particularly problematic is perception: the anti-realist may very well hold that what perception presents us with is shaped by our conceptual scheme or something such, wherefore what we perceive is reality-for-us, not reality-itself. But the very idea of *perception*, as opposed to mere *imagination*, requires that there *is* some kind of influ-

[40] See also, e.g., Goodman [1978].

[41] This is, very briefly, the complaint Eli Hirsch likes to make for some metaphysical disputes. However, he doesn't take this stance in every case. See, e.g., Hirsch [2002a,b, 2005, 2008].

[42] For the external/internal distinction, see Carnap [1950].

ence 'from outside', that is, from an independent reality—reality-itself (we return to this point in §1.2, see p. 35 below). The anti-realist thus holds that it is futile to try to characterize reality as it is independently of us, yet this reality-itself is still taken to *exist*: it is just that we cannot but comprehend it in the form of reality-for-us. It is problematic, however, to claim that *there is* this reality-in-itself, even though we cannot provide any positive characterization of it whatsoever. For what exactly is thus being claimed? It is reasonable to demand from any metaphysician that she be able to explain *what* exactly she takes there to be (a point we elaborate on in §2.2.1). That is, the metaphysician should be able to provide positive characterizations.[43] Yet as soon as the anti-realist starts providing such characterizations for reality-itself, she has moved towards realism: the provided characterizations are bound to be couched in some conceptual scheme or other, yet they are supposed to characterize mind-independent reality. And that is precisely what the realist does. It appears that the anti-realist is forced to admit some minimal, hidden portion of realism in order to be able to state her view in the first place.

Secondly, there is the thorny issue of what *we ourselves* are supposed to be. Either we take ourselves to be what we are (say, human beings, or persons, or rational agents) because of our conceptual scheme, or we don't. Suppose that we do: then we are to believe that 'the authors *in* the stories are the *real* authors' [Putnam 1977, p. 496], that is, we are to believe that we are both authors and protagonists of the same big world-story. When taken literally, this statement resembles the story of Münchhausen who pulled himself out of a swamp by grasping his own wig and lifting himself up. And if we are not to take it literally, well, then, as Elder puts it, the statement comes closer to Zen *koans* than to serious (analytic) philosophy [Elder 2011, p. 53]. So we should better not take this line: we are what we are independently of our conceptual scheme. Then, however, we have to conclude that we occupy a metaphysical position different from that occupied by all other things in the world, given that they are dependent on our conceptual schemes while we are not. This would commit the anti-realist to a view of ourselves as transcendental egos of sorts, which is quite controversial, to put it mildly—at least within analytic philosophy.

Thirdly, and finally, allowing the conceptual schemes that belong to us as subjects so fundamental an influence on the make-up of reality entails a deep and unsettling

[43] On some interpretations of Kant, he has tried to answer this demand by claiming that his notion of a thing in itself is a *Grenzbegriff*, a notion that lies at the very boundary of what makes sense for us, and hence only admits of negative usage—see, e.g., Kemp Smith [1918]. Strawson [1966] appears to adopt a similarly negative understanding of the notion—which is why he claims that 'Kant, as transcendental idealist, is closer to Berkeley than he acknowledges' [p. 6].

kind of relativism—even if it is ensured that we humans necessarily share the same basic conceptual scheme. For it comes down to the thought that our very aim to understand reality is misconceived: there is no reality to *discover*, only a reality to *make*—be it by convention and hence arbitrarily, or by the inevitable make-up of our minds and hence necessarily.

These three at least potentially problematic aspects of anti-realism have, together with further such aspects, been subject to continuing debate—we are not presenting anything new on this count.[44] We rehearse these specific worries because they, again, are symptomatic of the metametaphysical picture underlying the anti-realist position, and it is this underlying picture that interests us. We can now see that this metametaphysical picture appears to be, at root, the same as that underlying the realist's project: the Modern picture. There is a distinction between the world as it is in itself and the world as it is for us, in science and/or common sense. Reality as it is for us is thought to depend somehow on our own faculties of perception and conception. The anti-realist contends, in addition, that our conceptual scheme applies *only* to reality as it is for us (indeed, that it is partly constitutive of that reality), and hence thinks it futile, or even incoherent, to theorize about reality-in-itself. One can even find anti-realists reluctantly speak of this reality-in-itself as an 'amorphous lump' [Dummett 1973, p. 73] or as just 'stuff' [Sidelle 1989, n. 11 on p. 55, 1998, §VI]—determinations that are supposed to convey that no determinations are possible, since every determination is inevitably a *conceptual* determination and hence not mind-independent.[45] The most famous and explicit example of this kind of view is, of course, the Kantian characterization of reality-in-itself as containing *Dinge an sich* ('things in themselves'): the very idea of such a thing in itself is of something not determinable in a conceptual manner.[46]

The realist, determined as she is to uncover the true make-up of reality as it is *anyway*, complains that anti-realism amounts to a kind of skepticism, a categorical denial of knowledge of the independent, external world: we are trapped in *our* reality. And, as Stroud writes: 'The thought that we can have no knowledge of things as they are independent of us is what makes skepticism so distressing' [1984, p. 162]. Giving up on independence does not relieve the distress, so the realist feels—hence if the anti-realist's arguments against realism are sound, this would lead the dedicated

[44] For a recent discussion see, e.g., Elder [2011, ch. 2].

[45] See also Jubien [1993], and Eklund [2008] for a more recent defense of the 'amorphous lump'-picture. And see Blackson [1992] for a nice elaboration of the 'stuff'-picture which takes it to be a 'commitment to the possibility of alternative semantic conventions' [p. 67].

[46] We do not mean to subscribe to any particular interpretation of Kant in writing this down—the remark is merely meant to illustrate.

realist to adopt a form of metaphysical pessimism rather than anti-realism (we will discuss this option in more detail below, in §1.4).

As an aside, note that among anti-realists there is divergence in views too, of course, but it is far less pronounced than in the realist case, and hence does not lead to such impasses as can be observed in the realist discussion. Given many anti-realists' overtly pragmatic understanding of metaphysical theorizing, differences in views won't bother them as much as in the case of the realists: this, indeed, is thought to be a major virtue of the anti-realist standpoint. Anti-realists are mainly in the business of defending anti-realism against realism (and here we do have something like an impasse, too). In fact, versions of what we call metaphysical anti-realism often go by realism-resembling names, such as 'internal realism' [Putnam 1981, 1983] or 'quasi-realism' [Blackburn 1993, esp. Part I], which nicely illustrates the desire to get rid of unwanted relativist and idealist connotations by way of presenting the anti-realist view as being very close to realism but without the realist's problematic metaphysical commitments.

Sometimes, it is not even clear whether a professed 'realism' amounts to metaphysical realism or anti-realism in our sense. For instance, Ernest Sosa has argued, in his reflections on one of Putnam's versions of anti-realist realism, for a position that is sometimes called 'explosive realism':

> Reality is objectively much richer and more bizarre than is perhaps commonly recognized. All sorts of weird entities derive from any given level of particulars and properties. ... Our objective metaphysics is hence absolutist ... given our inability to find any well-motivated objective restriction on the matter-form pairs that constitute derived entities. Our relativism applies to the truth or falsity of ... ontologically committed claims. It is here that a restriction is imposed by the conceptual scheme of the claimant speaker or thinker. [Sosa 1999, p. 143]

Sosa, in effect, proposes to include *all* the objects to which any conceptual scheme could give rise in reality as it is in itself—that is what 'absolutist' means in the above quote. It is thus indeed a kind of explosion of reality, which Sosa here proposes as a viable 'objective metaphysics'. Our conceptual scheme then simply restricts our attention to those objects that correspond to it. It seems therefore that Sosa ends up being a metaphysical realist, albeit of a rather unusual stripe. Yet, an anti-realist could state very similar things while intending to convey just that *there is no objective restriction* as to which conceptual scheme can be correctly applied to reality. The difference between these readings is subtle but profound.[47]

[47] We will revisit Sosa's explosive realism in the context of our discussion of laws of nature in §5.2.2.

This concludes our survey of those features of realism and anti-realism that illustrate the underlying Modern picture. We should keep in mind that our survey has not aimed at completeness: we merely touched upon a number of interesting questions one may ask when considering metaphysical realism and anti-realism. Our aim was to sketch a general framework for distilling the underlying Modern picture. The next section systematically develops and critically reflects on this Modern picture, which we have thus discovered to underlie most of the contemporary metaphysical discussion on both sides of the realism/anti-realism divide.

1.2 Exploring the Modern Picture

In the previous section, we extracted a first approximation of the philosophical picture that forms the background for large parts of the current metaphysical discussion, for realists and anti-realists alike. As metaphysicians, we are concerned with reality as it is in itself, either to provide a positive theory as to what this reality is fundamentally like (realism), or to argue that no such theory is possible, that we can only inquire into reality as it is for us (anti-realism). Let us unpack this first approximation carefully.

It is crucial to both parties that we distinguish between reality as it is for us and reality as it is in itself: this distinction is required for the very formulation of the realist and anti-realist projects. We have been deliberately vague on what exactly one should take 'reality-itself' and 'reality-for-us' to mean, since that differs considerably amongst versions of realism and anti-realism. A nice way of bringing out what exactly we wish the distinction to capture is to indicate how it relates to Sellars's famous and colorful distinction between the manifest and the scientific image.

On the one hand, Sellars says, we have the manifest image, which pictures the world as we, on reflection, take it to be; as such it contains 'not only "cabbages and kings", but numbers and duties, possibilities and finger snaps, aesthetic experience and death' [Sellars 1962, p. 35]. On the other hand, there is the scientific image, which results from the scientific practice of postulating entities that explain the manifest world: it takes its point of departure from the manifest image, but goes on to construct an independent conception of reality, of how things *are*, as opposed to how we take them to be (in the manifest image). Now, the task of the philosopher, according to Sellars, is to bring those two images together in stereoscopic vision in such a way that no clash ensues (roughly). Since we're actually in the middle of developing the scientific image, this will involve (1) getting clear on what the scientific image may look like, (2) on what the manifest image consists in (consulting common sense), as

well as (3) uncovering problematic aspects of either image as one compares the two.[48]

Sellars's view on the task of philosophy illustrates a standpoint that the more scientifically-minded metaphysical realist might take: the scientific image is meant to capture reality-itself while the manifest image is merely our parochial, anthropocentric, subjective take on reality, or in other words, captures reality-for-us but not reality-itself.[49] Anti-realists, on the other hand, will urge that the scientific and the manifest image are just two alternative versions of reality-for-us, and that neither of them is any closer to reality-itself than the other, though the one or the other may be more useful in certain important respects.

What is crucial to the Modern picture is the very idea of a reality-itself that is in some difficult to grasp sense 'more real' than reality-for-us, where it is left open how we should conceive of this reality-itself. Realists are in the business of constructing a decent positive conception of it, usually out of certain privileged ingredients taken from reality-for-us, but, as we have seen, frequently also by simply 'advancing bold and surprising claims' (see §1.1.1). Anti-realists aim to convince us that it is impossible to construct a positive conception of reality-itself, and that we should therefore convert to some kind of conception of reality-for-us that we can or should accept on transcendental, social, conventional, or other non-realist grounds.

The first characteristic metametaphysical principle of the Modern picture thus reads:

> *Separation* Reality as it is in itself is to be strictly distinguished from reality as it is for us.

Reality as it is for us obviously involves us in some way. We gestured towards the way in which realists and anti-realists understand this involvement: on the one hand, reality-for-us is thought to involve us in the way it is perceptually present for us; on the other hand it is thought to involve us in the way we conceptualize it. How exactly this involvement is spelled out differs between realists and anti-realists, and also among realists and among anti-realists. But it is clear that *Separation* requires a certain take on how we, in our various everyday or scientific or philosophical investigations, relate to reality-itself and to reality-for us. In particular, a certain understanding of how we are confronted with reality in perception and of how we go about conceptualizing what we thus find before us is presupposed. We may

[48] See Sellars [1962], and see Stanford [2012] for an interesting critique of Sellars's images talk and his related fairly metaphysical understanding of philosophy.

[49] Other versions of realism will locate our scientific view of the world more on the side of reality-for-us, thus claiming the realm of reality-itself for metaphysics proper. Notice that Sellars would probably not approve of our way of describing the manifest image as 'merely' reality-for-us.

capture the required orientation in the following two principles, one epistemological, the other metaphysical:

Source	Perception consists in affectations of our senses by reality-itself, and hence constitutes our primary access to that reality.
Mind-Dependence	Our concepts are the mere products of our minds.[50]

Our formulation of these two principles is bound to raise some eye-brows. We therefore elaborate on both of them with some care.

Those familiar with the philosophy of perception know that especially the formulation of *Source* is sensitive to several controversies.[51] Yet for our purposes it is enough if we focus on the core idea, and not on the tricky details: perception puts us in connection with reality-itself, while our concepts do so only via, say, their extensions (on a realist reading), or not at all (on an anti-realist reading). What we aim to say is not that perception *presents* us with reality-itself, because this will generally be denied both by the realist and by the anti-realist. On both views, the connection to reality-itself is indirect: what we are *directly* confronted with, in our experience, is reality-for-us, not reality-itself (although some realists will go on to argue that at least certain aspects of our perceptual experience of reality do 'reach all the way' to reality-itself). What *Source* aims to capture is the thought that perception puts us in contact with reality-itself, even if what this contact results in is merely reality-for-us.

It might still be surprising that reality-itself plays such a central role in *Source*: isn't this supposed to be the starting point also for the anti-realist? Would not an anti-realist simply deny that perception puts us in contact with reality-itself? Such a denial would be highly problematic, because, as we observed earlier (see §1.1.2, p. 29), even the anti-realist will have to take our perceptual faculties to constitute a point of contact between reality-*itself* and us—indeed, this is one of the crucial points where the anti-realist position depends on the existence of such a reality-itself. The product of this interaction, our perceptual experience of the world, will, from an anti-realist point of view, be *incontrovertibly* shaped by peculiarities of our perceptual and cognitive make-up. Hence, as said, what we find, in experience, is reality-for-us, not reality-itself. If even this is denied, we end up with a view on which the experiencing subject somehow makes up her own experiences instead of receiving

[50] This principle is related, in interesting ways, to the nominalist position in the medieval debate on universals (see Pasnau [1997] for an interesting approach to that debate). It is, however, impractical to name this principle 'nominalism', because its proper counterpart would then be 'realism', which is bound to be very confusing. Furthermore, 'nominalism' already has quite some uses within analytic philosophy that are likely to confuse as well.

[51] A brief look at, e.g., the essays collected in Noë and E. Thompson [2002] makes this clear.

them from outside, in which case there would not really be any *perception* anymore. It is open to the anti-realist, of course, to deny that we can even *state* this dependence of reality-for-us on reality-itself, be it in our way or in any other way, but it should be kept in mind that the ground for denying the very statability of this dependence is, precisely, that the dependence holds. This is one of the cryptic and vexing aspects that are typical of the more sophisticated versions of anti-realism.[52]

We move on now to *Mind-Dependence*. This principle, saying that concepts are the products of our own minds, is the very opposite of the central thesis of this essay: that essences are concepts. But we should not get ahead of things to come in the next section. On the Modern picture, how exactly concepts relate to reality-itself is a delicate matter, both for the anti-realist and for the realist. The anti-realist takes reality-for-us to be populated by objects and events etc. that are partly constituted by our concepts; our concepts manage to refer to the thing(s) they designate (in a broad sense of 'thing') because these things owe their presence, in reality-for-us, to these concepts. The anti-realist thus owes us an explanation of how this 'constitution' works and what it involves. The realist, on the other hand, takes our concepts to be tools for capturing, *inter alia*, the metaphysical structure of reality-itself (roughly), and hence faces the task of explaining both what such 'metaphysical structure' (or whatever alternative notion is preferred) consists in and how our concepts manage to mirror that structure by referring to aspects of it.[53] For our purposes, it does not matter much how exactly these or related questions are answered. Rather, *Mind-Dependence* is meant to capture just the need for such answers. In fact, our formulation of *Mind-Dependence* is too strong in this respect, because we don't wish to exclude views on which concepts are abstract objects of sorts, and hence not products of our minds. The crucial point is that concepts are not (or at least not in general) constitutive of their instances, that is, of mind-independent, external reality-itself.

Separation, *Source* and *Mind-Dependence* together form the backbone of the Modern

[52] If the anti-realist rejects such dependence of reality-for-us on reality-itself in *every* way (empirical, transcendental, semantic, implicit, conceptual, or whatever), he is in danger of losing grip on his own position: why should the reality he does acknowledge still have to be reality-*for-us*, and not simply *reality*? The very idea of a reality-for-us requires the contrasting notion of reality-itself. If he replies that there is *only* reality-for-us because it is our own construction, then we should point out that a construction requires *some* prior reality—the reality of whatever is doing the constructing. If that is then said to be also just part of reality-for-us, incoherence threatens. We start wondering what motivated the search for such a problematic position in the first place—a good motivation for considering an alternative to the Modern picture.

[53] This is, very roughly, Rorty's 'sentence-shaped chunks'-objection against realism [Rorty 1989, p. 5]: reality, Rorty thinks, does not speak a language, hence does not contain ready-made chunks to make true our sentences—at best, the world can cause us to believe a certain sentence true (although the truth of this statement about causing beliefs is itself already problematic for him).

picture, the basic metametaphysical picture informing the metaphysical discussion. On its basis, we can characterize realism and anti-realism, together with the third, skeptical option we have occasionally mentioned above, as follows:

Skepticism Given that we can only get to the real world with the help of our own concepts (*Mind-Dependence*) via our own perceptual inputs (*Source*) it is impossible to say whether anything we take ourselves to know about the real world is really true. Behind the veil of sensory inputs, to which we apply our concepts, a reality radically different from our conception of it might exist. We can never know.

Anti-Realism Given that every claim to knowledge is phrased by use of our own concepts or language (*Mind-Dependence*) as applied to what our senses deliver (*Source*), we must concede that we can only claim knowledge of the world as it is *for us*, and not of the world as it is *in itself*. *Skepticism* and *Realism* have to be rejected because they involve a mistaken understanding of what the target of our cognition is.

Realism *Skepticism* and *Anti-Realism* have to be rejected because they don't do justice to our scientific and/or commonsense understanding of and interaction with reality. Using our own concepts (*Mind-Dependence*), which are always subject to refinement and enrichment through scientific progress, we are able to reconstruct what reality in itself, to which we only have indirect access via our perceptual input (*Source*), is like.

The skeptical option has not been of concern for us up until now, because within the contemporary discussion it constitutes more of a challenge rather than a position actually defended. However, there is a less radical version of it, which we have mentioned earlier under the label 'metaphysical pessimism'. Such metaphysical pessimism does not conclude that we can never know reality-itself, but rather takes our realist quest for understanding reality to be so hard that progress may be possible only to a very limited extent (we discuss this kind of view, and its relation with skepticism, in §1.4 below).

To illustrate the relations between these three options, it is helpful to understand them as responding in different ways to the following two philosophical issues that arise from *Separation*, *Source* and *Mind-Dependence*:

Justification We need to justify our claims to knowledge of the world as it is in itself, given that what we have is merely (1) how that world presents itself to us through our senses (*Source*) and (2) our own concepts (*Mind-Dependence*).

Truth We have to make sense of how any judgment we make with the help of our own concepts can be true of reality in itself, which may involve making sense of how our concepts/words can be properly connected to things in reality-itself (reference).

The skeptic's challenging answer to *Justification* is that there is no way for us to justify any claim to knowledge about the world as it is in itself (and the metaphysical pessimist concludes that there are at least grave if not insurmountable obstacles). The implicit understanding of truth for the skeptic is one of correspondence to how things are in reality as it is in itself—truth does not involve any epistemic component.

The anti-realist agrees with the skeptic's answer to *Justification*, but adds that we in fact only need to justify claims about reality as it is for us. The accompanying explication of truth naturally involves an epistemic component: to claim that some judgment is true is to claim that it is justified or warranted, in some appropriately idealized sense. In effect, as we have seen, the anti-realist claims that the world we perceive and think about is reality as it is for us, not the reality as it is in itself—it is the former that should (and does) concern us, not the latter. And since reality-for-us depends on our epistemic make-up, in some sense, what is true within reality-for-us similarly involves an epistemic aspect.

The realist, however, argues that we *can* justify claims to knowledge of reality as it is in itself, by appealing to an inference to the best explanation: reality as it is in itself should explain the way it presents itself to us (i.e., reality-for-us). The accompanying explication of truth as correspondence agrees with that implicit in the skeptic's position. A paradigm contemporary realist explication of truth would be in terms of *truthmaking*: our true judgments are made true by the world's being a certain way (e.g., by certain states of affairs or facts obtaining, or by certain objects' instantiating certain properties or standing in certain relations, or merely by the objects themselves that thus instantiate certain properties or stand in certain relations, etc.).[54] The nice thing about truthmaking, for the realist, is that he can plug in his favored metaphysics to yield appropriate truth-makers—talk about oak trees, say, can be made true by space-time worms, or by features of the one blobject, or by relations amongst certain real numbers, etc. (cf. §1.1.1 above).

Let us put our findings with regard to the relations between skepticism, realism and anti-realism into a little scheme:

[54] See, e.g., Armstrong [2004b] and the essays in Beebee and Dodd [2005] and Lowe and Rami [2008].

	Skepticism	*Anti-Realism*	*Realism*
Can we arrive at knowledge of reality-in-itself?	No	No	Yes
Does truth involve correspondence with reality-in-itself?	Yes	No	Yes

One way to understand the realism/anti-realism dichotomy is, then, by taking both the realist and the anti-realist to be responding to the skeptical challenge: the latter by rejecting the skeptic's understanding of truth, the former by rejecting the skeptic's take on justification.[55]

We now have a more substantial sketch of the metametaphysical picture we call the *Modern picture*, and of how it gives rise to the realism/anti-realism dichotomy. We found out, in the previous section, that it indeed stands in the background of the contemporary (meta)metaphysical debate. Adopting the Modern picture results, as we have indicated, in an oscillation between two problematic positions: finding skepticism unsatisfactory, we can either become realists but be at a loss as to how we can adjudicate between highly divergent realist metaphysical theories, or we can admit our alleged inability to discover the nature of reality as it is in itself and make do with the anti-realist restriction of our knowledge to reality as it is for us—which in turn quickly leads to serious issues, such as a commitment to the idea of a reality-itself that is denied any content, that again motivates a return to realism.

As we observed at the beginning in this chapter, philosophers like John McDowell, Hilary Putnam, David Wiggins, Fiona Ellis, Sebastian Rödl and Michael Thompson have, from various angles, tried to overcome this uneasy situation by, in effect, questioning the very metametaphysical picture on which it depends: the Modern picture.[56] The present essay in general, and this first chapter in particular, aims to contribute to this trend by exploring an alternative to the Modern picture specifically from the point of view of (contemporary analytic) metaphysics.

From a metametaphysical point of view, we may thus ask what kind of picture emerges once we reject the problematic Modern picture, which we characterized by formulating the principles *Separation*, *Source*, and *Mind-Dependence*. Rejection of these principles gives rise to an alternative metametaphysical picture. The following

[55] For those who incline towards systematicity: the fourth position, which answers the justification question positively but the correspondence question negatively, is rather odd when understood along the lines of the Modern picture. However, on the Aristotelian picture we wish to defend in this essay, it is in fact a very natural view. See fn. 65 on p. 44.

[56] See fn. 15 on p. 17 above.

section aims to formulate, in very general outline, such an alternative to the Modern picture, which we develop and defend at greater length in chapters 2 and 3. And from a metaphysical point of view, we may then go on to ask to what kind of metaphysical view this alternative picture gives rise. The second part of this essay is devoted to a survey of such a metaphysical view.

1.3 The Aristotelian Picture

Before we engage in metaphysical thinking, we tend to take things as they present themselves: we are naive realists of sorts. And that is what we are, to a large extent, in ordinary life too, even after we started doing metaphysics. From this naive point of view, we see ourselves as subjects who are directly confronted with objective reality through our faculties of perception. As soon as we start to critically examine this naive, unreflective picture, we move towards the Modern picture. We observe, for example, the crucial role that our sense organs play for our view on reality, the ease with which they can be deceived, the fact that various animals are sensitive to features of the environment we are not sensitive to (and *vice versa*): in short, we are pushed towards a subjectivist understanding of certain perceptible features of things. On the other hand, we notice that our concepts play an important role, too: my ecologist friend tends to see a lot more during our walk in the forest than I, because I lack many of the required concepts. We notice considerable differences in conceptual schemes amongst different cultures. We come to recognize how intricate the connection between our conceptual framework and our view on the world is, both in our understanding and in our perception of it. In short, we come to suspect that what we naively took to be reality is in fact merely reality *as it is for us*, and that we may be far removed from reality *as it is in itself*. We have come to replace the naive, unreflective picture with the Modern picture; in its wake, skeptical doubts about our ability to gain any knowledge at all of reality as it is in itself arise.

We have identified *Separation* as the root of the Modern picture. It installs a gap between reality-itself and reality-for-us, and once we have that gap, we have to find a way to deal with it. The problems of *Justification* and *Truth*, in their typical form, become urgent: we are pulled in the opposite directions of realism and anti-realism.

Now that our philosophical eyes have been opened, we cannot go back to the naive, pre-reflective view. Yet we have entertained considerations that put the Modern picture in an unfavorable light: it leads to a duality of unsatisfactory metaphysical

viewpoints—realism on the one hand, anti-realism on the other hand. Might we not have made a mistake in adopting the Modern picture? We start to wonder whether it is possible to reject it without thereby giving up on critical reflection on our world view. The present section aims to sketch our alternative to the Modern picture in very rough outline. We call our alternative picture the Aristotelian picture. It is, as it were, 'post-skeptical'.

It is important to note that rejecting *Separation* does *not* imply rejecting the innocuous distinction between how things seem to be and how things really are. That distinction was already present within the naive picture, after all.[57] Rather, rejecting *Separation* means discarding the picture according to which reality-itself is fundamentally hidden from us, and reaches us at best merely indirectly, via reality-for-us, which we experience and live in. What this rejection implies for our relation to reality is best illustrated by first revisiting the other two principles constitutive of the Modern picture, *Source* and *Mind-Dependence*.

As an alternative to *Source*, we may consider McDowell's conception of experience as *openness to the world* (as developed in his [2011]). To use his typical wording, the idea is that in experience we 'take in how things are', instead of being confronted with the mere effects reality-itself exerts upon us, which we then have to relate to a reality hidden behind them.[58] This approach to perception is usually called 'direct realism' in the philosophy of perception. McDowell's way of fleshing out this idea comes down to the proposal that what we take in, in experience, is not devoid of concepts, but rather already endowed with conceptual content.[59] The exact details of this proposal do not matter for present purposes—only two of its central tenets: what we receive through perception in its many forms is *not* the mere effect of a hidden reality-itself, and it is *not* disjoint from the conceptual.

This brings us to the other principle, *Mind-Dependence*. The central claim of this essay is, in fact, the alternative to *Mind-Dependence* that we need at this point. This alternative is easy to state, but harder to make decent sense of: essences are concepts.

[57] It is common wisdom that our senses can be unreliable, of course. See, e.g., Descartes's famous remarks on this score—on the towers that look round from a distance but square from closer by—in the *Meditations on First Philosophy* [Descartes 1641, p. 53, in the Sixth Meditation] (or, in the standard Adam and Tannery edition of Descartes's collected works [Descartes 1964–76], see volume 7, p. 77).

[58] See McDowell [1994a, p. 25].

[59] On such a view, problems related to hallucination and illusion are probably best dealt with by adopting some version of disjunctivism. Disjunctivism is the view that episodes of hallucination and illusion are different from episodes of veridical perception, wherefore there is no need to assume an intermediary layer, a 'veil of ideas', to be present in both kinds of cases. Disjunctivists are thus opposed to such a 'veil of ideas', which constitutes an extra layer between reality-itself and the mind, and thereby gives rise to a view on perception that is congenial to the Modern picture. See, e.g., the essays in Haddock and Macpherson [2008].

A first, straightforward attempt to make sense of this claim might run as follows: concepts are *constitutive* of their instances, rather than being mere pointers to their extensions, mere products of our own minds. Consider, for instance, *sortal concepts*. If we take a sortal concept to have a certain content that is, ideally, explicable in some kind of definition that captures at least the conditions of identity, persistence, and existence for its instances, we could take that definition to be the 'real definition' of the things to which the sortal concept applies.[60] A similarly essentialist understanding of properties could also be developed, for example along the lines of Elder [2011, ch. 8]. Elder defends a view according to which properties occupy a certain position within a range of contrary properties. A paradigm example would be, e.g., *having a mass of 10kg*, which contrasts with *having a mass of 11kg* etc. in such a way that all contrasting properties exclude one another. The essence of a property, we could say, consists at least in the range to which it belongs as well as the position it occupies within that range. Presumably, the essence of a property includes more; we could for instance include interrelations between different ranges—e.g., whatever has a certain mass also has a certain size and shape and location, etc.[61]

It is useful to compare this first gesture towards an essentialist understanding of concepts with realist and anti-realist approaches to concepts. The metaphysical realist typically takes concepts to be mental or abstract objects to which extensions correspond in reality-itself. The concepts are mere pointers to these extensions. Perhaps they embody parts of our knowledge of such objects, or useful rules for object classification that fix their extensions, but they are not part of the metaphysical make-up of the objects themselves. Put in Fregean terms: the 'level of sense' is carefully distinguished from the 'level of reference', and the former is taken to be metaphysically irrelevant for the latter. The anti-realist, on the other hand, takes objects to be located in reality-for-us, and to be partly constituted by our conceptual framework. For the anti-realist, an object is always an object *as* picked out by a certain sortal concept; and that's precisely why it cannot be an inhabitant of reality-itself.

Now, on the alternative picture we are developing, there is no reality-itself versus reality-for-us distinction. Yet we are attempting an understanding of the material things inhabiting reality more or less as the anti-realist would have it, i.e., as partially constituted by concepts, but without the anti-realist's mind-dependent understand-

[60] We will develop the thoughts presented here much more thoroughly in chapter 2.

[61] In §3.3.2 below we provide further motivation for considering such an essentialist understanding of properties, as opposed to an understanding based purely on extensions. In the second part of this essay we will sketch a metaphysical view that recognizes an even more differentiated understanding of properties.

ing of those concepts. That is, the things in reality are still as mind-independent as the realist could ever demand, simply because their essences are independent from whatever we may think of them. It turns out, then, that without *Separation*, we may combine these nice features of anti-realism and realism—viz., concepts as partially constitutive of objects and mind-independence of those objects, respectively—without running into the problems we have sketched earlier.

The rest of this essay is devoted to developing and defending the foundations of a metaphysical view on the basis of the Aristotelian picture we are outlining here. Chapters 2 and 3 are devoted to spelling out and defending the claim that essences are concepts—the former focusing on essences, the latter on concepts. In Part II we then move on to metaphysics proper. There, we take our cue from the work of Michael Thompson, especially from his [2008]. Thompson carefully develops a series of distinctions between ways concepts figure in our judgments concerning fundamentally different kinds of things—abstract objects, physical things, living beings (we introduce these distinctions in much more detail in chapter 4).[62] That is to say, Thompson at no point defends our claim that essences are concepts (nor does he reject the claim), but, as soon as we add in that claim, his view on these different types of conceiving of objects, of things, *ipso facto* amounts to a suitable essentialism. What we are working towards, then, is a differentiated view on how essences operate, that is read off from the way our actual thinking about these various realms of nature works (without, of course, claiming infallibility with respect to these diverse ways of thinking—see also §1.4 below).

Here we come to recognize what makes rejecting *Mind-Dependence* so difficult: the adherent of the Aristotelian picture must explain how concepts can be operative, find expression, in real, concrete phenomena of various kinds. Put in Aristotelian terms, the task is to come up with a proper notion of *formal causation*: a concept/essence becoming manifest in concrete phenomena, such as an electron behaving just as an electron should in certain circumstances, or a tree protruding branches just as that particular kind of tree should.[63]

At this point, the purpose of our sketchy remarks on essentialism is not to articulate a clear-cut version of it, but rather to illustrate the role the idea of essence can play in the context of a non-Modern picture. Interestingly, when reading contemporary

[62] Thompson's discussion focuses on action theory and ethics. The fact that he finds himself forced to develop his own metaphysics along the way shows to what extent metaphysicians are under the spell of the Modern picture nowadays.

[63] We give content to the notion of formal causation we have in mind in §4.3, p. 139; in §6.1.4, p. 199; and in §7.2.2, p. 270.

essentialists it is not always clear whether they operate against the background of the Modern picture, and thus conceive of themselves as metaphysical realists (in our sense), or are rather to be understood as contributing to the project we are interested in here, i.e., the project of providing an alternative to that picture.[64] However, the very idea of essentialism is, as will become clearer in chapter 2, not something over which one can disagree: even Modern picture-based metaphysicians have to accept essences in some form or other (real essences for the realist, conventional essences for the anti-realist). Of course, the essentialist metaphysical realist will have to add that he takes our concepts not to *be* essences but rather to *correspond* to essences. Thereby, he is saddled with the difficult task of showing what such 'non-conceptual' essences look like. For the anti-realist, such a task is self-defeating: any attempt to give content to the idea of such a non-conceptual essence is, after all, bound to be couched in conceptual terms. These complications do not arise on the Aristotelian picture, which does not install such a gap between the conceptual and the real, after all.[65]

Up until now, we have completely ignored what can appear to be a serious drawback of our essentialist understanding of concepts: we seem to have the ability to cook up concepts as we wish, and these are fairly often not even beginning to be adequate (as the history of both science and philosophy shows). The situation is rather similar to that in the case of perception: we want room for mis-conception much like we want room for mis-perception, but still allow successful conceptual activity to reach all the way to the world. Now, one way to make sense of this is by construing our faculty of conception to yield more or less complete *graspings* of concepts—in a word, *conceptions*—which may succeed or may not succeed in capturing the *concepts* operative, as essences, in reality.[66] That is, although we are capable of grasping the wrong concepts (or no concepts at all), we are also capable of grasping the right ones. Obviously, this requires careful unpacking, which starts

[64] The work of E.J. Lowe illustrates this difficulty nicely: to some extent, his metaphysical picture fits into the Aristotelian picture we are sketching here; but on the other hand several aspects of his philosophy seem to commit him to the Modern picture. See, e.g., Lowe [2008a, 2009]. The same holds for Kit Fine's work—see, e.g., Fine [1994b, 2001, 2003, 2005c]. Good examples of essentialists that quite explicitly reject the Modern picture in favor of an Aristotelian alternative are Wiggins [2001] and Oderberg [2007].

[65] As an aside, notice that the correspondence theory of truth will have to be rejected on similar grounds: true thoughts (which involve concepts) are not true because they *correspond* to facts, which then have to be conceived of in non-conceptual terms (*per impossibile*). It seems, then, that truths are better construed as *identical* with facts, just like concepts are identical with essences. That is why, in terms of the scheme we presented above (see p. 39), the Aristotelian picture *accepts* that we can justify knowledge of reality but *denies* that truth consists in correspondence to reality-itself. An in-depth discussion of the nature of truth on the Aristotelian picture, however, falls outside of the scope of the present essay.

[66] Here we draw on David Wiggins's useful distinction between conceptions and concepts; see his [2001, pp. 8–11].

by taking our faculty of conception to share certain features with our faculty of perception.[67]

Does such a distinction between concepts and conceptions tacitly reintroduce a Modern picture-like separation between reality as it is for us (conceptions) and reality as it is in itself (concepts)? No, because, in successful cases, conceptions simply *are* concepts. We have done nothing to eliminate the possibility of systematic and widespread error, both on the conceptual and on the perceptual side—but that is not our aim here (and we believe it not to be a very interesting aim in the first place: it arises out of the skepticism on which the Modern picture is based). The important point is that our proposed alternative to *Mind-Dependence* does not locate the conceptual realm within the confines of our human minds, or in some abstract realm disjoint from the external world of concrete objects, but instead grants the conceptual a fundamental metaphysical role in that concrete realm, independently of our minds.

Such a mind-independent, essentialist understanding of the conceptual meshes nicely with the understanding of perception as openness we borrowed from McDowell: we do not take in how things are *simpliciter*, but only insofar as we have grasped the right concepts, insofar as we have worked our way to incorporating the right conceptions into our conceptual repertoire. Whereas perception is in an important sense a passive capacity, through which reality presents itself to us, conception is an active one: if we don't engage in thinking about what we perceive, the flow of experience just passes us by. Hence the 'grasping' metaphor (which is, of course, itself difficult to grasp).

Let us now formulate these two alternatives as principles contrasting those embodied in the Modern picture, *Source* and *Mind-Dependence*, to which they are supposed to provide an alternative:

[67] In fn. 59 on p. 41 above, we mentioned that our principle *Openness* is likely to include a form of disjunctivism: cases of perception are, on such an account, different from cases of misperception. In a similar vein, one may envisage a conceptual form of disjunctivism: in cases in which one's putative conception does not correspond to any concept at all, one is simply not having a conception in the first place—just as hallucinations are simply not perceptions. We could even argue from perceptual disjunctivism to conceptual disjunctivism—for instance by drawing on Gareth Evans's remarks on demonstrative thoughts: on his view, one may take oneself to have a demonstrative thought, e.g., 'This apple is red', while hallucinating the apple, so that there is no object to demonstratively refer to. In such a case, however, one is not *really* entertaining a demonstrative thought, since there is no such thought to entertain—demonstrative thoughts are 'Russellian', that is, they require the existence of their objects. See Evans [1982, ch. 6, esp. p. 173]. Generalizing, there may be other cases (besides demonstrative thought) in which one takes oneself to be grasping a concept while failing to do so (perhaps familiar examples like phlogiston are cases in point).

Openness In our experience of the world we are in direct contact with reality.[68]

Mind-Independence The concepts which we use to think truly about the world are part of that very world (and not mere products of our minds).[69]

These two principles do not require a separation between reality as it is for us and reality as it is in itself, as *Separation* has it. However, a different separation is still required: a separation of reality in its perceptible and its conceivable aspects. Like the alleged separation between reality-for-us and reality-itself, this separation has everything to do with our organization as epistemic subjects aiming at knowledge. The picture is roughly as follows: we have access to perceptible aspects of things, the way they empirically present themselves, via our perceptual capacities (which is not to say that we have access to *all* perceptible aspects of things, of course), while we have access to their conceptual aspects, to what they are (their essences), via our faculty of conception. By combining these two aspects in the right way we arrive at reality. In short, perception teaches us *that* it is, conception *what* it is.

The crucial difference between *Separation* and the separation we have on our alternative, Aristotelian picture is that the latter separation can be overcome through the process of combination just sketched, whereas the former separation is by definition irreconcilable: reality-for-us contains ingredients that reality-itself *cannot* inhabit, viz., concepts as constitutive of things. The proper alternative to *Separation* for our Aristotelian picture is thus the following:

Combination By combining the perceptual and conceptual aspects of reality in the right way, we put together what in reality is never separated.

Herewith, we have arrived at a systematic first delineation of the *Aristotelian picture*, which we may set against the Modern picture we examined in the previous section as an alternative.

Returning to our observation that metaphysical realism leads to an ever more diverging range of views, we may ask whether the Aristotelian picture will prevent such divergence to occur. The Aristotelian picture leaves open many interesting

[68] Compare Aristotle in *On the Soul*: 'what has the power of sensation is potentially like what the perceived object is actually; that is, while at the beginning of the process of its being acted upon the two interacting factors are dissimilar, at the end the one acted upon is assimilated to the other and is identical in quality with it' [1984, p. 31], II.5, 418a3–6.

[69] Again, this principle corresponds in interesting ways to the realist position within the medieval debate on universals—just as its counterpart, *Mind-Dependence*, corresponds to the nominalist position. See fn. 50 on p. 35 above.

and important issues, of course, on which rival views can be developed—that much is quite trivial. The important point is that the divergence we noted within the realist discussion has a quite specific source: realism's problematic requirement of providing an account of reality-itself while at the same time urging that reality-itself is both disjoint from reality-for-us and utterly devoid of concepts (in the sense we have explained). This is what made it so difficult to make decent sense of the theoretical virtue of 'fit', as we have seen (see §1.1.1, esp. p. 26). There is no such tension within the Aristotelian picture, hence no reason to expect the realist's specific divergence problems to recur.

1.4 Metaphysical Pessimism

Up until now, we have motivated our investigation into an alternative to the Modern picture on the basis of our observation that, on that Modern picture, the metaphysical project faces certain typical problems—with regard to the fundamental opposition between realism and anti-realism, but also internally, within the realist and anti-realist projects themselves. It is interesting to see what happens if one shares our reservations concerning the realist and anti-realist projects, but not our search for an alternative to the Modern picture. We have occasionally mentioned such a view, and labeled it metaphysical pessimism—because it looks like this view leads to a rather pessimistic assessment of the prospects of metaphysical inquiry. We have put this position in the vicinity of the view we labeled *Skepticism* earlier (see especially §1.2, p. 38 above).

In his recent, thought-provoking discussion of the prospects of metaphysical inquiry, which proceeds in very fundamental terms but is supported by a range of carefully conducted case studies on the metaphysics of color (see Stroud [2000]) and on causation, necessity and value (see Stroud [2011, esp. and resp. chs. 2, 3, and 4]), Barry Stroud indeed moves towards such a pessimistic position. It is instructive to see where exactly his views diverge from ours, and why, as it helps make clear our own position considerably.

Stroud argues that metaphysical results, or, as he prefers to say, negative and positive 'metaphysical verdicts' concerning something that is regarded metaphysically problematic, are much more difficult to arrive at than is usually thought.[70] He complains that:

[70] What Stroud calls negative metaphysical verdicts is what we have earlier called local anti-realisms. See our brief remark on p. 19 above.

> in works of philosophy it is usually all over by the bottom of page one. I think that . . . can be confirmed by more or less random reading. What really matters is off the page and settled in the mind before the author's announced task has even begun. [Stroud 2000, p. ix]

One of the reasons Stroud has for arriving at such a rather sweeping diagnosis of contemporary philosophy—and of metaphysics in particular—is that, on reflection, virtually all arguments in support of a negative metaphysical verdict rest upon prior endorsement of a certain picture of reality that already implies that negative metaphysical verdict. This is what is 'settled in the mind' before the metaphysical argumentation starts. For example, concerning color Stroud writes:

> Prior acceptance of the exclusively scientific story of the physical world is what encouraged the idea of perceptions of color as nothing more than 'sensations'. [ibid, p. 182]

So far, Stroud seems to be largely on our side. Metaphysical realists who wish to defend metaphysical views of reality-itself that are much more austere than reality-for-us, and hence have to endorse quite some negative metaphysical verdicts, tend to settle 'what really matters', which is the Modern picture plus their preferred conception of reality-itself, in their minds before their metaphysical project is even begun.

Another reason for Stroud to be ambivalent about negative metaphysical verdicts is the observation that many philosophers propound such verdicts on one or another issue—e.g., color, causality, necessity, value—whilst not realizing the impact that such a verdict would have, if *really* endorsed with all its consequences, on our overall world view. To illustrate what Stroud has in mind here, consider the belief in witches during the early modern period in Europe. Many people were convinced that there are witches—hence the horrible witch hunts and burnings that took place until deep into the 17th century. Gradually, people came to endorse a 'negative metaphysical verdict' concerning witches and witchcraft. The result was, of course, that people *revised* their overall world view accordingly: they did no longer see signs of witch-involvement, they no longer suspected people of being witches, they no longer believed that there is such a thing as witchcraft. In short, they eliminated the whole family of witch-related concepts from their considered conception of reality. One might consider this to be a successful case of arriving at a 'negative metaphysical verdict' concerning witchcraft.

Now consider color. Most philosophers of perception believe that color is in a certain important sense not real. But it is not so easy to really amend our world

view accordingly. We cannot help but believe that lemons are yellow just as firmly as that they have a typical shape—even if we go on to proclaim, in our philosophy classes, that their yellowness involves something subjective which their shape does not.[71] This observation shows that it is indeed not as easy to arrive at negative metaphysical verdicts concerning such basic aspects of our conception of reality as many philosophers think: apparently, our world view strongly resists the accompanying revision. Still, these philosophers continue to hold such subjectivist views on color. This indicates involvement of the Modern picture. The color subjectivist thinks it unproblematic that we continue to endorse our ordinary beliefs about color, because those concern reality-for-us whereas the 'real' story on which color is 'nothing more than sensations' concerns reality-itself. The tension Stroud uncovers then disappears—and with it an important constraining factor for metaphysical theorizing, as we said (see §1.1.1 above).

Kit Fine even goes so far as to suggest that *because* negative metaphysical verdicts such as that about color can be endorsed without giving up on ordinary color beliefs, there has to be a distinction between reality-for-us and reality-itself:

> Is there room for [negative metaphysical verdicts] ... that does not put them in conflict with received opinion? If there is, then it requires that we be able consistently to affirm that something is the case and yet deny that it is really the case. It requires, in other words, a *metaphysical* conception of reality, one that enables us to distinguish, within the sphere of what is the case, between what is really the case and what is only apparently the case. ... Truth is one thing, metaphysical status another. [Fine 2001, pp. 2–3]

Fine contrasts this sophisticated kind of negative metaphysical verdict with a more straightforward type of local anti-realism that consists in the 'mere rejection of what we ordinarily accept'. He thinks, however, that:

> many of us are inclined to doubt that philosophy is in possession of arguments that might genuinely serve to undermine what we ordinarily believe. [ibid, p. 2]

We sharply disagree with Fine here: by resorting to the Modern picture, he indeed secures space for negative metaphysical verdicts that will not serve to undermine what we ordinarily believe, but the costs are high, as we have seen. As our example of witchcraft indicates, we do leave room for negative metaphysical verdicts even on the Aristotelian picture, but only insofar as they really make a difference, that is, insofar as they *do* 'serve to undermine what we ordinarily believe'.

But let us return to Stroud. Things get more problematic than in the color case if we move to more fundamental topics, such as causation. In the case of color, there at

[71] See Stroud [2000, esp. pp. 178–80].

least appears to be room for an argument to the effect that we cannot help but believe things to be colored, despite their not really being colored in the way we believe they are, because of the way they affect our senses. Stroud calls such strategies 'unmasking' strategies: color is purportedly unmasked to be 'nothing but' the effect certain structural features or dispositions of things have upon our perceptual make-up (say). A negative metaphysical verdict concerning causation, on the other hand, cannot be arrived at by such an unmasking strategy, since the unmasking itself has to proceed in causal terms:

> Our causal beliefs are to be 'unmasked' as due to something about us and not to any causal connections that hold in the independent world. ... We even use that same causal vocabulary in understanding how the world led us to acquire and understand those very terms in the first place. [Stroud 2011, p. 54]

There seem to be limits, then, as to what features from reality-for-us one can defend a negative metaphysical verdict about.

From here on, Stroud moves towards a view very different from ours. He arrives at a rather pessimistic overall picture: metaphysical inquiry, he concludes, is a very serious undertaking indeed, yet the very conceptual presuppositions of that undertaking stand, at least in some fundamental cases, such as causation, in the way of arriving at any positive or negative metaphysical verdicts at all. Nevertheless, he urges, we should continue with our 'quest for reality', for it might just be that by working really hard we will be able to arrive at a positive or negative verdict after all, in some case or other—and if not, we will at least have deepened our understanding of the metaphysical project and it limits.

Stroud's pessimism is based, in large part, on a kind of transcendental argument. To see how this argument goes, we need to take a brief look at what the metaphysical project embodies, according to Stroud. To engage in metaphysical inquiry in the first place, he holds, is no trivial or innocent matter. Consider the following passages:

> Metaphysics must proceed 'from within.' We must start from the conception of the world we already have ... [ibid, p. 145]
>
> [A]ny conception of the world we ... try to reach metaphysical conclusions about will include human beings and their experiences and thoughts and beliefs and actions. ... It is not something obvious and easily open to view. [ibid, p. 146]

We need a certain starting point for metaphysical inquiry, which, Stroud claims, can only be our considered world view so far—what we would call, on the Modern picture, reality as it is for us. With regard to the goal of the metaphysical enterprise, Stroud writes:

> The goal is to achieve an enhanced—a metaphysically corrected—conception of what the world is like. [Stroud 2011, p. 8]

That is, we conceive of the metaphysical project as moving towards a metaphysically purified conception of reality. The project proceeds by critically examining certain aspects of our considered world view, which may involve eliminating these aspects, or reconstruing them in metaphysically less problematic terms, or perhaps revising them in some other way.

The project of thus critically examining certain aspects of our world view proceeds by means of impartial examination:

> It is an attempt to stand back and not take for granted those parts of our conception of the world that are for the moment under metaphysical scrutiny. [ibid, p. 5]

It requires, thus, that we take a certain neutral stance towards the aspects of our world view under consideration: endorsing them and rejecting them should both be live options, and the outcome should depend on what our metaphysical investigations yield.

Now, in the case of causation, as we saw, Stroud argues that we cannot eliminate the idea of causal dependence from our overall world view because it is involved in explaining how we come to have a world view at all. Hence causality is in an interesting way *indispensable*: we cannot take the required neutral stance, hence we cannot possibly arrive at a negative metaphysical verdict. The point is indeed a transcendental one: such a metaphysical outcome is ruled out by the very possibility of engaging in metaphysics. Now, Stroud is not as optimistic as Kant was about such transcendental observations—the indispensability of causality, he argues, implies that *we* have to think of reality in causal terms, but not that reality *itself* has to involve relations of causal dependence. Indeed, the very indispensability of causality stands in the way of our ever arriving at a *positive* metaphysical verdict as well. Our metaphysical reflection simply does not get off the ground: we cannot help but endorse the kinds of causal beliefs that we have. Hence no metaphysical vindication of those beliefs can be achieved. The best we can do is attempt to understand better why we are unable to arrive at either metaphysical verdict. Perhaps we will find out that we were wrong about the indispensability of causality, but even if we don't—as Stroud thinks—we will still learn something, namely what the limits of our metaphysical quest for reality are.

All in all, then, Stroud paints a picture of the metaphysical project which is very different from ours, which indeed results in a rather pessimistic estimate as to what

we can expect the metaphysical quest for reality to yield. We already conjectured that this is what happens if one takes to heart the problems that plague the metaphysical project without rejecting the Modern picture, within which this project is framed. Let us now see whether this diagnosis indeed applies to Stroud's views.

Consider, first, the following passages:

> Metaphysics seeks more than a description of our conception of the world. To believe something ... is not necessarily to take a stand on any metaphysical question. [Stroud 2011, p. 141]
>
> Being unable to say unambiguously what the metaphysical question asks ... can encourage the thought that there is no meaningful metaphysical 'question' to 'answer.' ... [But] there is a metaphysical urge, or need, that cannot be denied, however difficult it is to express it in unambiguous terms.. . . What is remarkable is the strength of the widespread feeling that things simply must be metaphysically one way or the other. [ibid, p. 159]

Stroud makes a sharp distinction between our considered, overall conception of reality and a metaphysical conception of it; our 'ordinary' beliefs about reality do not all by themselves amount to 'metaphysical' beliefs about reality—even if the 'ordinary' beliefs have been thoroughly examined in the light of (scientific or everyday) evidence. Coming to metaphysical conclusions, Stroud urges, requires a specifically metaphysical quest, and hard though it is to state the aim of that quest, it is a quest we really do engage in—like ourselves, Stroud opposes both an anti-realist denial of the quest as well as the kind of quietism that seeks to free us of the quest by showing it to be based on confusion.[72]

Looking a bit more closely now at Stroud's conception of the metaphysical project, we find some striking remarks:

> Metaphysical reflection tends to concentrate on how such things as causal dependence ... are to be accommodated within, or otherwise somehow excluded from, *what we already take to be so* in the independent world. [ibid, p. 152; emphasis added]
>
> I want to *start out with a minimal conception of the world* that seems about as "absolute" as any human conception could be, and see what must be added to it to eventually arrive at as full and adequate a conception as we need. [Stroud 2000, p. 45; emphasis added]
>
> A positive metaphysical verdict would have been reached if ... the very contents of the things we believe in the areas in question can be reduced without remainder

[72] Such quietism, famously held by Wittgenstein [1953], can be found in the work of, e.g., the later Putnam [1991, 1992], Putnam and Conant [1994], and Putnam [1999], but also in McDowell [1989, 1994a, 2009b,c]. However, McDowell's take on 'philosophical activity conceived, in the somewhat Wittgensteinian way I favour, as "therapeutic" ' [McDowell and Willaschek 2000, pp. 100–1] does not appear to be negative in the way Stroud opposes: by removing misconceived philosophical problems and views, McDowell does not end up with a rejection of the quest for reality but rather with a positive view on what that quest involves.

> to *something that is uncontroversially so* in the world as it is independently of us and our responses to it. [Stroud 2011, p. 124; emphasis added]

Stroud seems to suggest, in these remarks, that there is some kind of partition of our considered world view into (1) those aspects that are 'uncontroversial' in the sense that they embody 'what we already take to be so'; and (2) those aspects that fall outside this 'minimal conception' of reality. The 'uncontroversial' aspects of our conception of reality thus *do* have a certain metaphysical status, despite their being 'ordinary' beliefs. This partition makes sense against the background of the Modern picture: we start out with a conception of what is in fact reality-for-us, and are, as metaphysicians, in the business of transforming it into a conception of reality-itself. Stroud is much more careful than most metaphysical realists in restricting the ways in which we can arrive at a suitable conception of reality-itself, as we have seen, but still, the underlying metametaphysical orientation appears to be the same. Stroud, we may conclude, comes very close to being a metaphysical realist in our sense—and hence is sailing under the flag of the Modern picture.

It is interesting, in this regard, that Stroud is almost exclusively concerned with negative metaphysical verdicts, barely with positive ones. The few things he does say about such positive verdicts are, in fact, quite telling—consider, for example, the following statement:

> A positive metaphysical verdict [about causation] must differ in some way from everything we believe that is expressed in causal . . . terms. It seems possible to hold beliefs of those kinds with no metaphysical opinions one way or the other. What then would a positive metaphysical verdict add to what is already expressed in those beliefs? [ibid, p. 158]
>
> Accepting a positive verdict, along with the irreducibility of the beliefs, would seem to leave us with [causal] facts as simply otherwise inexplicable aspects of reality. And if that is how they seemed at the beginning, we would remain dissatisfied, having made no metaphysical advance. [ibid, p. 159]

Here, again, Stroud asserts that believing something is not all by itself a metaphysical commitment. But there is also a certain expectation of what a metaphysically purified conception of reality should deliver: something much more austere than reality-for-us, in terms of which reality-for-us should be explained. This structure is familiar from our discussion of metaphysical realism.

Yet it is unfitting to simply put Stroud on the list of metaphysical realists, because of the noted ambivalence towards realist metaphysics—or, rather, towards negative metaphysical verdicts. We do better, indeed, to understand him as moving towards *Skepticism* (see §1.2, p. 38 above): the skeptic, on our conception, shares the realist's

understanding of what the metaphysical project is about, i.e., arriving at a conception of reality-itself, but finds compelling reasons for thinking that no progress is possible, that metaphysical disappointment is inevitable. Now, Stroud is obviously not an across-the-board skeptic in this sense—in fact, he writes, in conclusion of his case study into the metaphysics of color:

> I do not say it can be proved once and for all that such disappointment is inevitable in the quest for the reality of the colours of things. Any convincing proof of such a verdict would give us finality and so a kind of metaphysical satisfaction after all, and so it could not be sound. [Stroud 2000, p. 191]

That is, Stroud encourages us to continue our quest for reality. However, his conclusions with regard to *indispensable* aspects of our conception of reality, such as causation, necessity and value, can be put under the heading of skepticism: in effect, he claims that (1) there is a real question as to whether there is any causation, necessity and value in reality, and (2) that our putative knowledge of these topics cannot be metaphysically justified (see §1.2).

Stroud at one point remarks that '[t]he question is about the relation between the conception we have of the world and the world itself' [Stroud 2011, p. 6]; it is precisely this question which leads him, in the face of indispensable aspects of our conception of the world, to his pessimistic—or, as we may now say, skeptical—conclusions about the limits of the metaphysical project. And it is precisely this question on which, as we have seen, the Modern picture and the Aristotelian picture markedly diverge.

Put in Stroudian terminology, the Aristotelian picture-based approach to metaphysics that we are developing here accepts that our current, considered conception of reality forms the starting point of metaphysics, but rejects the distinction between 'ordinary' and 'metaphysical' beliefs as well as the assumed partition of the former into those beliefs that we regard to be metaphysically 'unproblematic' and those that we do not. In effect, then, we reject that there is a specifically metaphysical quest for reality, distinct from ordinary or scientific investigation. On our view, what we aim to understand is reality; the metaphysical project only differs from ordinary and scientific investigation in non-essential ways, e.g., in its broadness of scope—usually, the concepts metaphysics cares about (such as causation) are simply assumed in those other kinds of investigation. That is why it makes sense to inquire into, e.g., the *metaphysical* presuppositions and implications of contemporary physics or biology.[73]

More specifically, however, the Aristotelian picture takes the way reality hangs together to simply *be* conceptual—essences are concepts, as our main thesis has

[73] We will illustrate such inquiry in the second part of this essay.

it. Stroud, however, finds this step, from the conceptual to the real, to be highly problematic. He denounces

> the familiar but hollow arrogance of idealism: things must be a certain way because we thinkers must think things are that way. But without idealism, how can we make the further step to reality? [Stroud 2011, p. 143]

We argue that it is a mistake, grounded in the Modern picture, to think that there is a step to be made here. Does that imply that the 'familiar but hollow arrogance of idealism' applies to us? We would think not—at no point did we argue that because *we have* to think in certain ways, *reality itself* has to be that way. Rather, we say this: we find that certain concepts apply. Hence those concepts structure reality, unless we are mistaken. Whether or not some of these concepts are indispensable, in Stroud's sense, is neither here nor there. It might be slightly unsettling to learn that *if* we are wrong about the applicability of such indispensable concepts, there is no way for us to find out nor to adjust our views accordingly. But the indispensability is not our *ground* for believing reality to obey them.

We have provided some reasons for thinking that Stroud, who is operating against a broadly Kantian background, is wedded to the Modern picture. Yet he rejects anti-realism as simply giving up on a 'quest for reality' which he judges to be a serious matter, while also noticing the problems that arise once the 'quest for reality' is formulated as the metaphysical realist has it. The result is that he is forced in the direction of skepticism—hence his overall pessimism.

We offer an alternative: abandonment of the Modern picture in favor of the Aristotelian picture. We do not provide arguments that show with apodictic certainty that the Aristotelian picture is right while the Modern picture is wrong. We merely point towards the fact that adopting the Modern picture is a choice we have made during the development of our scientific and philosophical understanding of reality over the last centuries, and that there is reason to reconsider that choice. The proof of the pudding, we would like to suggest, is in the eating—which means that we need to explore our proposed alternative to such an extent that we may start to experience some actual eating.

Our alternative of course offers no response to the skeptic: it does not give any answer as to how one might be *certain beyond any doubt* that one has grasped the right concepts. Doubt is always possible—but in some cases simply irrelevant. When in doubt, we can mount further investigations. When our doubt takes the form of global skeptical scenarios, for which no further investigations seems possible, we may set it aside as idle speculation.

1.5 Concluding Remarks

We have uncovered and criticized the Modern picture, which lies behind the principal divide between metaphysical realism and metaphysical anti-realism. This Modern picture, we showed, is based on a fundamental distinction between reality-itself and reality-for-us (*Separation*), and includes a certain picture of how we, with our faculties of perception and conception, relate to this two-layered reality. Perceptually, we are in contact with reality-itself, but what it shows is reality-for-us (*Source*). Conceptually, we start out with reality-for-us and, according to the realist, end up with a conceptual reconstruction of reality-itself, or, as the anti-realist has it, can never break out of reality-for-us (*Mind-Dependence*). We saw that Barry Stroud, though very critical and cautious about both approaches, still ends up implicitly endorsing the Modern picture, resulting in his pessimistic appraisal of the prospects of metaphysical inquiry.

We motivated the search for an alternative to this Modern picture by discussing several characteristic features of realism and anti-realism, and of the kind of metaphysical pessimism Stroud arrives at. We outlined an alternative, Aristotelian picture, which is based on the denial of *Separation*, *Source* and *Mind-Dependence*, to yield *Combination*, *Openness*, and *Mind-Independence*. On this picture, both perception and conception put us, when correctly combined, in direct contact with reality, the one presenting to us *that* it is, the other telling us *what* it is.

The Aristotelian picture, not unlike the Modern picture, comes with challenges of its own, some of which we have already acquainted ourselves with—to repeat, it is quite impossible to sketch something as general as the Modern or the Aristotelian picture without already going into details that call for a much more thorough philosophical examination. Yet, the challenges that the Aristotelian picture faces are very different from the problems to which the Modern picture gives rise, and hence it is likely that the menu of options available for metaphysics against the background of the Aristotelian picture is different from the menu of options described earlier, too. As already noted, a very prominent task for the proponent of the Aristotelian picture lies in making sense of *Mind-Independence*, the idea that the concepts we use in thinking about the world somehow *belong* to that world, instead of just being our own creations which merely *correspond* to an utterly non-conceptual reality-itself. This essay is devoted to precisely that task—we have gestured towards an essentialist take on this issue which we will develop further in the remainder of this essay.

It should be noted, as a reminder, that this task has a counterpart within the Modern picture as well: to make sense of the idea of a reality that is entirely *devoid* of

concepts. The anti-realist holds that this is one of the main reasons why we cannot coherently state any positive view on reality-itself, yet she herself is still committed to the existence of such a non-conceptual reality-itself, as we have seen. The realist, on the other hand, faces the difficult task of showing how our conceptual knowledge of reality-itself can be true of it even though reality itself is utterly non-conceptual.

The Aristotelian picture, however, promises not to lead us to an irresolvable oscillation between unsatisfactory views, like the Modern picture does. It offers a way out of the pessimistic impasse to which this unfortunate situation can give rise. Whether or not the Aristotelian picture can keep this promise depends on whether we have correctly diagnosed the root of the realism/anti-realism dichotomy within the Modern picture: *Separation*, the assumption that there is a fundamental gap between reality as it is in itself (which is non-conceptual) and reality as it is for us (which is conceptual). For, once we have the gap, we want to close it, but we cannot really close it without giving up *Separation*—and that amounts to rejecting the Modern picture.

Part II of our essay is devoted to a sketch of the metaphysical outlook to which the Aristotelian picture gives rise. At the same time, it aims to clarify what is at stake in certain key contemporary metaphysical debates—on laws of nature, on causation, on time, and on life. One lesson we will draw is that the Modern picture encourages many metaphysical realists to isolate *part* of our conception of reality and claim that *only* that part corresponds to reality-itself. That is a form of metaphysical dogmatism—apparently justified by a skeptical attitude towards the other parts of our conception of reality. It gives rise to reductive views concerning everything that falls outside of the isolated conceptual fragment. The tendency to think along these lines turns out to be very strong indeed, even amongst those resisting reductionism in one or another form. The Aristotelian picture leaves no room for it, for on that picture there is no gap to be bridged between the conceptual and the real.

Combination, *Mind-Independence* and *Openness* provide the basis for a metametaphysical picture that differs fundamentally from the Modern picture, on virtually all philosophically interesting dimensions. It does not lead to the kind of Wittgensteinian quietism one may find in the work of McDowell, the later Putnam, and others[74]: we take the picture to invite a philosophical research program of its own, one which aims to discuss the issues we concern ourselves with here, and many more. Indeed, our essay is devoted both to developing foundational aspects of this research program and to developing a first approximation as to what a metaphysical view based on the Aristotelian picture looks like.

[74] See fn. 52 on p. 52.

Chapter 2

Essential Truths*

It is our aim both in this chapter and in the next chapter to defend and elucidate our claim that essences are concepts. In this chapter, we focus on the essences-side of that claim, so to speak, whereas in the following chapter we will be concerned with the concepts-side. In particular, the present chapter is devoted to developing the foundations for a suitable version of essentialism—suitable, that is, for the Aristotelian picture as outlined in the previous chapter (see §1.3). We do not, however, defend any particular essentialist claims in this chapter; we stay on the metametaphysical or methodological level. Indeed, an important result of our investigations below will be that essentialism is, properly understood, *not* a metaphysical doctrine—not even for those working under the Modern picture. It is, rather, constitutive of metaphysics as such (and thus belongs under the heading of metametaphysics). We will descend from this meta-level and get involved in a number of substantial metaphysical issues in Part II of this essay (in chapter 7, for instance, we will defend a form of biological essentialism).

Our point of departure is the peculiar inferential step one frequently finds in the literature from principles of individuation[75] to essential properties of the things individuated. For instance: sets are individuated by their members, hence sets have their members essentially. We argue that this inference is indeed valid, given a certain understanding of what principles of individuation are, and how they relate to the broader project of the metaphysician. In particular, we conclude that the essential

*Large portions of this chapter have been drawn from Mulder [2013].

[75] In the literature, what we call 'principles of individuation' are also often referred to as *criteria of identity*. Notice, though, that there are diverging conceptions of both available. See Lowe [2003] for discussion.

features of things are those that are implied by the *sortal concept* they ultimately fall under. Such an ultimate sortal concept is best characterized as *ontologically fundamental*. This notion of ontological fundamentality is one to be carefully developed below (in §2.2.1). Putting these results into the context of our Aristotelian picture, we arrive at the view that the ontologically fundamental concepts identified by the metaphysician, if he is successful, coincide with the concepts that structure reality. E.g., assuming *set* to be an ontologically fundamental sortal concept, it is that very concept which embodies the relevant principle of individuation, hence it is *that concept* which grounds the essentialist truths. Grasping that concept amounts to grasping the real essence of sets: essences are concepts.

Our focused investigation into the sketched essentialist inference results in a defense of essentialism in abstraction from specific essentialist claims. Such claims, based on applications of the essentialist inference, appear widely in the metaphysical literature. Saul Kripke, for instance, has famously argued for the essentiality of origin for physical objects on the grounds that such objects are individuated (partly) by their origins.[76] And the neo-Aristotelian line taken by, e.g., David Wiggins, David Oderberg, and E.J. Lowe[77], who take objects to be individuated by the (substantial or ultimate) sortal under which they fall, illustrates a somewhat more complicated implementation of the essentialist inference, one that is in fact close to the understanding of essentialism to be defended below—the sortal, in their view, provides a principle of individuation which determines synchronic as well as diachronic identity for its instances. Interestingly, virtually the same claims have been defended by anti-realists such as Alan Sidelle and Amie Thomasson, who wholeheartedly agree with our yet to be established diagnosis that the kind of modality involved in the essentialist inference is conceptual modality, but add that concept choice is a matter of convention so that the modality is merely conventional and not real.[78] (See the previous chapter for discussion of the realism/anti-realism issue, and the next chapter for our understanding of conceptual modality.)

Of course, it has turned out to be notoriously difficult to articulate principles of individuation for concrete objects. Sets (and other abstract objects) are much easier, in that respect, and for this reason, we focus for some part of our discussion on the example of sets, although our conclusions apply generally.

Below, we start with an exploration of the essentialist inference in §2.1: what is

[76] See, e.g., Kripke [1972, pp. 112–115]: '[A]nything coming from a different origin would not be this object' (p. 113); he understands origination properties to be 'essential properties' [p. 115].

[77] See, e.g., Wiggins [1995, 2001], Oderberg [2007], and Lowe [2006, 2009].

[78] See Sidelle [1989, 1998, 2010] and Thomasson [2007].

the typical form such an inference takes? Can we understand it to be a formal inference, on a traditional understanding of 'formal'? These first, probing considerations provide the basis for answering, in §2.2, the critical question whether the essentialist inference is indeed valid, and if so, why. What undergirds the essentialist inference, we argue, is the specific job principles of individuation are supposed to perform within the larger project of the metaphysician, which we characterize on a highly abstract level. It is here that we develop our notion of ontological fundamentality, which underlies the essentialist inference but also underlies the broader essentialist view that any conceptual truth associated with the fundamental ontological category a given thing instantiates grounds essential truths about that thing (in the broadest possible sense of 'thing').

We continue to sketch, in §2.3, the implications of our construal of essentialism, and of the essentialist inference in particular, by reference to some concrete, realist metaphysical views. In the final section of this chapter, §2.4, we tie our results somewhat firmer to the broader context of this essay.

2.1 Exploring the Essentialist Inference

The essentialist inference takes us from principles of individuation to essentialist conclusions. In order to prevent confusion, it is helpful first to clarify what is at stake in a preliminary fashion (in §2.1.1 below). Having thus cleared our way, we continue to present a formal rendering of the essentialist inference, in order to locate where exactly the putative essentialist conclusion comes from (in §2.1.2). We then argue that a purely formal-logical approach to the essentialist inference is unfruitful, mainly because whether or not some statement is a principle of individuation is not merely a matter of its logical form (§2.1.3). The lesson learned in this first, exploratory section is thus mainly negative, yet it does point us in the right direction for our further investigations in §2.2 below.

2.1.1 Preliminaries

Before delving into the precise workings of the essentialist inference, a few things need to be clarified. The first concerns the distinction between epistemic and metaphysical principles of individuation, the second is the difference between essentiality and necessity, and the third is the difference between general and individual essences.

In this chapter, we are concerned with ontological principles of individuation, not

with epistemological such principles. The distinction can be put as follows: while epistemological principles of individuation capture how *we* individuate things, or perhaps provide heuristics for settling identity questions, ontological principles of individuation purport to capture how things *themselves* are individuated. One quite important difference between the two notions is that there can be very many different ways *for us* to individuate things, while ontologically speaking this turns out to be rather problematic. More content will be given to the idea of such ontological individuation in due course (especially in §2.2.1 below); for now it is enough to flag the distinction.

As Fine [1994b] has convincingly argued, necessity and essentiality are not the same. It is necessarily true that Socrates is a member of the singleton {Socrates}, but, while it *is* essential to {Socrates} that Socrates is a member of it, it is *not* essential to Socrates that he is a member of {Socrates}. Fine glosses 'essentially true' as 'true in virtue of the essence of . . . ', where the blank needs to be filled in by some object (or class of objects). This explains the noticed asymmetry. The modality involved in essentialist statements thus differs from 'ordinary' metaphysical modality. Yet if a statement is true in virtue of some essence (in Fine's sense), then it cannot be false, and hence will be necessarily true as well—so there is a close connection between the two notions. What the relations between these types of modality precisely are is not our concern in this chapter; it will be in the next (see especially §3.3.3 and §3.4). As indicated, we will eventually identify this special type of modality with conceptual modality.

Essences come in various guises. Some people identify an object's essence with the (set of) properties it essentially has[79], while others, most prominently neo-Aristotelians, hold that a thing has its essential properties *in virtue of* its essence. Fine's conception of essence is a nice case in point.[80] We argue, of course, in favor of our thesis that essences are concepts, which implies that the latter construal of essence is to be preferred. Here are two typical statements of that construal:

> Just as we may define a word, or say what it means, so we may define an object, or say what it is. [Fine 1994b, p. 2]
>
> To define something just means, literally, to set forth its limits in such a way that one can distinguish it from all other things of a different kind. . . . Putting the point again in Aristotelian terminology . . . , to give the definition of something it to say *what it is*, to give the *ti esti* or *to ti ēn einai* of the object. Put simply, the real essentialist position is that it is possible to say correctly what things are. [Oderberg 2007, p. 19]

[79] See, e.g., Plantinga [1974, esp. chs. IV and V], Yablo [1987], and Elder [2011, ch. 3, esp. §3.5].
[80] But see also, e.g., Wiggins [2001, chs. 3–4], Oderberg [2007, chs. 1–3], and Lowe [2009, chs. 2–3].

The thought is that the essence of some thing is the real definition of that thing, so that its having the essential properties it has follows from that definition.[81] We will slowly work towards this form of essentialism in the present chapter.

A further distinction can be made between *individual* essences and *general* ones. If sets have their members essentially, then *this* set, {Socrates}, has Socrates as a member essentially. Thus, having Socrates as a member follows from the individual essence of {Socrates}, but not from the general essence, considered in abstraction from this instance. Such a notion of individual essence does not always accompany defenses of general essence. For instance, Putnam [1975] can be understood to have argued that elm trees could not possibly fail to be elm trees, that they are essentially elm trees. All elm trees would, on this construal, have the same (general) essence. Principles of individuation, being principles governing identity and difference, are bound to result in individual essences (as our example of sets illustrates)—we will thus defend individual essences as well as general essences.

There is a possible confusion here. First, individual essences as intended here are not to be identified with the haecceities or primitive this-nesses involved in the bare particularist's view. Individual essences as involved in the essentialist inference are determinates falling under determinables, where the latter are specified precisely by the principles of individuation (and hence by the concepts) in question. Individual essences are, then, instances of general essences. As we will see (in §2.3), bare particulars are a limiting instance of this general idea.[82]

In what follows, we compare the modality to which the essentialist inference gives rise to conceptual modality, and in particular, as said, to the kind of modality involved in what is true 'by definition' (cf. §2.1.3). Insofar as such terminology gives rise to the question whether it should be read along realist or anti-realist lines, we can refer to the previous chapter: we primarily intend our investigations to elaborate the metametaphysical outlook embodied in the Aristotelian picture. We will, however, have occasion to remark on how our results bear on the Modern picture-based metaphysical realist and anti-realist standpoints as well.

Since our main example will be essential set membership, it is fitting to briefly remark on the one person (to the best of our knowledge) who has explicitly *denied* set membership to be essential: A.J. Ayer. Here is his argument:

> The pair set consisting of the Crystal Palace and the Eiffel Tower is also the pair set consisting of the largest building erected for the Great Exhibition in London

[81] The notion of 'following from' involved will be our topic in the next chapter.

[82] We will articulate a somewhat more detailed understanding of individual essences and their relation to general essences in §2.4 below.

> in 1851 and the largest building erected for the Great Exhibition in Paris in 1889. But it is obviously not necessary that these descriptions should be satisfied by the Crystal Palace and the Eiffel Tower. There is therefore a sense in which the pair set consisting of the Crystal Palace and the Eiffel Tower might not have contained either of them. The sense is that the set can be uniquely identified by a description which neither of them necessarily satisfies. [Ayer 1979, p. 312]

Surely the sense in which Ayer is right is merely the sense in which it is true that *to that description* there might have corresponded a different set, not that the set *itself* might have contained different members. We circumvent such semantic subtleties by defaulting to a *de re* reading of set membership attributions.

2.1.2 A Formal Approach

Let us carefully examine the workings of the essentialist inference by developing a formal representation of it. Since the essentialist inference is thought to apply to principles of individuation, we do well to start with a formalization of such principles. Now, a principle of individuation for *F*s typically takes the following form:

(PI) x and y are *F*s $\rightarrow (x = y \leftrightarrow \varphi(x, y))$

$\varphi(x, y)$ should, of course, amount at least to an equivalence relation over the *F*s, and it should be compatible with Leibniz's Law, which states that identicals share all their properties:

(LL) $x = y \rightarrow \forall\Psi(\Psi x \leftrightarrow \Psi y)$

The φ-part of a principle of individuation is, of course, the crucial part: it functions as the *individuator*. Presupposed is that whatever is involved in φ applies to *F*s—e.g., in the case of sets, the membership relation. One form which this φ in (PI) plausibly takes is that of a biconditional, possibly prefixed by a number of quantifiers, as follows:

(PI′) x and y are *F*s $\rightarrow (x = y \leftrightarrow \forall z_1, \ldots, z_n(\varphi'(x, z_1, \ldots, z_n) \leftrightarrow \varphi'(y, z_1, \ldots, z_n)))$

The essentialist inference now seems to work as follows. Suppose we have a principle of individuation for *F*s of the form (PI′), and suppose that x is an instance of *F*. If we now assume some instance of the φ'-part of that principle to be true of that x:

(EI1) $\varphi'(x, \mathsf{c}_1, \ldots, \mathsf{c}_n)$

we may derive the corresponding essentialist claim:

(EI2) $\Box_x \varphi'(x, \mathsf{c}_1, \ldots, \mathsf{c}_n)$

where the symbol $\Box_x$ stands for the essentialist modality (in Fine's [1994b] words: it is true in virtue of the nature of x that ...). Given this result, we may go on to conclude the universal closure:

(EI3) $\forall z_1, \ldots, z_n(\varphi'(x, z_1, \ldots, z_n) \rightarrow \Box_x \varphi'(x, z_1, \ldots, z_n))$

Applying this general scheme to our example of sets, we get the following. Here is the so-called Axiom of Extensionality, which functions as the principle of individuation for sets:

(Set) x and y are sets $\rightarrow (x = y \leftrightarrow \forall z(z \in x \leftrightarrow z \in y))$

The essentialist inference now licenses the inferential step from (S1) to (S2) and (S3):

(S1) $\mathsf{c} \in x$

(S2) $\Box_x(\mathsf{c} \in x)$

(S3) $\forall z(z \in x \rightarrow \Box_x(z \in x))$

This indeed gives us essential membership for sets.

Now that we have a rough conception of how the essentialist inference works, formally speaking, we may ask whether the essentialist inference can indeed be conceived of as an inference based purely on the form (PI) or (PI'), so that from any true statement of that form some essentialist conclusion follows. The answer turns out to be decidedly negative.

2.1.3 Why a Formal Approach Fails

A simple consideration immediately makes clear why the essentialist inference cannot apply to just any statement of the form (PI) or (PI'). For suppose that there is a creature who likes Rembrandt's paintings very much—call him *Rem*. For every painting Rembrandt has produced, Rem devotes precisely one entire day to contemplating just that painting. So now we may individuate Rembrandt's paintings by the day on which they get contemplated by Rem ($R(x, y)$ denotes the relation of Rem contemplating x on day y):

(Rem) x and y are paintings of Rembrandt $\rightarrow (x = y \leftrightarrow \forall z(R(x, z) \leftrightarrow R(y, z)))$

Now, it is obviously wrong to suppose that Rembrandt's paintings *essentially* get contemplated on specific days by Rem. There just happens to be, *ex hypothesi*, a suitable

injective map from paintings of Rembrandt to days, so that one can write down a principle like (Rem) that is actually true. But Rem might have had a different hobby, of course—he might have been fond of Shakespeare's plays instead of Rembrandt's paintings.[83] It seems, then, that (Rem) is just a (contingent) truth about paintings of Rembrandt, not a principle of individuation for them (although, if true, (Rem) could of course be useful as an *epistemological* principle of individuation—see §2.1.1 above).

Thus, not every statement of the form (PI) or (PI') expresses a principle of individuation from which essentialist conclusions can be drawn. But the essentialist inference is thought to apply precisely to principles of individuation. Hence, if our formal presentation of the essentialist inference is on the right lines, the essentialist inference cannot be based purely on the form (PI) or (PI').[84]

Let us consider a more serious example from the literature, in order to further illustrate what a principle of individuation is good for, but also to disentangle the project of *establishing* principles of individuation from the more general question as to how such principles work. Lowe and Williamson have had a discussion[85] on principles of individuation which centered on Frege's famous principle of individuation for directions ($d(x)$ is a function from lines to their directions and $\mathrm{Par}(x, y)$ denotes the relation of parallelism):

(Dir) $\forall x \forall y (d(x) = d(y) \leftrightarrow \mathrm{Par}(x, y))$

(Dir) states that the direction of two lines x and y are identical just if x and y are parallel. This principle is not of the form (PI), and Williamson [1991] argues that all attempts to reformulate it so as to fit (PI) are unacceptable. For consider the following two proposals (where $\mathrm{Of}(x, y)$ denotes the relation that holds between a direction x and a line y just if x is the direction of y)[86]:

(Dir') x and y are directions $\rightarrow (x = y \leftrightarrow \forall z_1 \forall z_2 (\mathrm{Par}(z_1, z_2) \rightarrow (\mathrm{Of}(x, z_1) \leftrightarrow \mathrm{Of}(y, z_2)))$

[83] Geach [1980] contains lots of interesting examples of such spurious principles of individuation—introduced by him in support of his thesis of relative identity, which allows for the following possibility: that a and b are the same F while not being the same G. Though it would be interesting to compare our discussion with his relativism, such an excursion would take us too far afield. Our own understanding of principles of individuation, to be developed below, is opposed to relative identity. For what it is worth, we agree with the arguments against relative identity put forward by Wiggins [2001, chs. 1 and 6] and Lowe [2009, ch. 5].

[84] This has already been observed by van Cleve [1985]. He uses a more abstract example, the identity of indiscernibles:

(II) x and y are things $\rightarrow (x = y \leftrightarrow \forall \Psi (\Psi x \leftrightarrow \Psi y))$

This principle has the form (PI'), but surely accepting (II) would not commit one to the extreme thesis that everything has all its properties essentially.

[85] See Lowe [1991] and Williamson [1991].

[86] The proposals we consider are different from the ones Lowe and Williamson consider; this is because, for reasons of clarity, we wish to stay as close as possible to (PI'), not just to (PI).

(Dir'') x and y are directions $\rightarrow (x = y \leftrightarrow \forall z(\mathrm{Of}(x, z) \leftrightarrow \mathrm{Of}(y, z)))$

The principle (Dir') seems faithful to Frege's idea to have parallelism play a major part in the individuation of directions, and hence looks like a suitable candidate principle of individuation for them. However, there seems to be no sound reason to prefer (Dir') over (Dir''), which works just as well and is simpler, but does not mention parallelism at all, and hence fails to capture the gist of Frege's idea. Now, we do not want two different principles of individuation for one and the same kind of thing: that would lead, via the essentialist inference, to diverging ascriptions of essential features—e.g., is parallelism essential to directions or not? So it seems that we have gotten ourselves into an awkward situation.

We should be careful, however, to distinguish Frege's project from the issue of principles of individuation in general. (Dir'') works just like (Set): it says that directions are identical just if they 'have' all and only the same lines. Whether or not one accepts this as *the* principle of individuation for directions depends on one's further metaphysical outlook—on one's take on abstract objects in general, and on one's views in the case at hand in particular. For instance, one may indeed follow Frege in associating objects with any partitioning induced by any equivalence relation, and thus accept directions to be *objects*; or one might argue that directions are determinate *properties* falling under the determinable 'has a direction', much like specific colors relate to 'has a color'.[87] Moreover, one might disagree with Frege about what is more fundamental and use (Dir'') to define parallelism in terms of directions, instead of *vice versa*. However that may be, there is one important lesson to draw, which we here take from Lowe:

> [I]t is surely not at all surprising that a criterion of identity [i.e., a principle of individuation] for *F*s, taken in conjunction with one or more further necessary truths concerning *F*s or related objects, should entail a proposition which, although not *itself* a criterion of identity for *F*s, states a necessary and sufficient condition for *F*-identity. [Lowe 1991, p. 193].

We don't need to worry about (Dir'') following from (Dir') together with certain further truths about lines and directions: if we have reasons to prefer (Dir') as the correct principle of individuation for directions, the truth of (Dir'') is simply irrelevant. The point is, again, that being a principle of individuation is not a purely formal matter.[88]

[87] We touch upon this understanding of properties—that is, as falling in ranges of contraries which mutually exclude each other—in §1.3, p. 42, and in §3.3.2.

[88] Failure to notice this fact leads to a trivialization of the very notion of a principle of individuation, as is exemplified by Noonan [2009], who alludes to examples that no one would seriously propose as principles

We argue that the essentialist inference does not depend solely on the form of principles of individuation on the basis of *just one* way of construing that form, viz., (PI'). Strictly speaking, that observation does not provide enough evidence for our claim—for all we have said so far, there may be other, more successful ways of construing the logical form of principles of individuation that do allow for such a formal reconstruction to be successful. Some reflection, however, provides the missing warrant. A principle of individuation will at the very least have to express a necessary and sufficient condition for identity, and hence will have to take the generic form $x = y \leftrightarrow$ ——. Now, however one fills in the blank, the problem is, as we have shown by way of example, that there is no guarantee that there are no *other* truths that happen to have that very same form.

In fact, there will *always* be such truths: one can derive from a given principle of identity, however one settles the question as to their canonical form, indefinitely many other truths of that same form by simply adding necessary truths, like $2+2 = 4$, to the individuator. Here is an example derived from (Dir''):

(Dir''') x, y are directions $\rightarrow (x = y \leftrightarrow \forall z((\mathrm{Of}(x,z) \wedge 2+2 = 4) \leftrightarrow (\mathrm{Of}(y,z) \wedge 2+2 = 4)))$

Applying the essentialist inference to all such derived truths would yield inclusion of all necessary truths in the essence of everything that has an essence.[89] This short consideration might all by itself be enough to conclude that the essentialist inference is not a purely formal matter, unless one is prepared to accept such inclusion of necessary truths as just an innocent or trivial aspect of essences. That is why we opened this section with a stronger argument that turns on inclusion of *contingent* truths into the essence of things (like, e.g., (Rem)).[90]

This concludes our first exploration of the essentialist inference. We learn that the essentialist conclusion follows from a statement having logical form (PI) only if it is being put forward *as* a principle of individuation—otherwise, we run into trouble. But what does it mean to put forward principles 'as' principles of individuation?[91]

of individuation (i.e., examples like the one we sketched above, involving Rembrandt's paintings) in order to argue, in effect, that such principles are much less informative than many people think.

[89] Thanks to an anonymous referee for the paper Mulder [2013] for pointing this out.

[90] In the next chapter we will develop a broader notion of form which extends beyond what is traditionally called logical form, on which the essentialist inference can be judged to be formal after all. The form on which the essentialist inference is based, on that alternative, in effect includes the fact *that it is* a principle of individuation, hence counterexamples such as (Rem) and (Dir''') are simply ruled out: there cannot be any statements sharing that form which are not principles of individuation. *That* form, however, is not statable using traditional logical symbolism (see §3.3).

[91] Michael Jubien takes this question to reveal 'an important obscurity in the concept of *identity criterionhood*' [Jubien 1996, p. 347]. He argues that the 'obscurity' is fatal, and concludes that identity is a primitive, 'kind-transcendent' relation [p. 351]. We discuss his skepticism about individuation below, in §2.2.1.

What is such a principle of individuation supposed to do? On a first approximation, it should settle questions as to whether two given *F*s are identical or not. And in order for it to provide such guidance, it should not *presuppose* the individuation of the very *F*s at stake.

It appears that we can read statements of the form (PI) in two importantly different ways. On the one hand, we can read them *stipulatively*, as principles of individuation that do not presuppose properly individuated things; on the other hand, we can read them *factively*, as just truths about things which we already have before us, properly individuated. We have presupposed a factive reading of (Rem) in our example above, and a stipulative reading of (Set). If we were to read (Set) as just a factive statement about sets, not a stipulative principle of individuation for them, it would *presuppose* a suitable realm of properly individuated entities, sets, about which things can be said—e.g., that they happen to be identical if and only if they are coextensive. On the other hand, if we read it stipulatively, as a principle of individuation of sets, it partly *constitutes* that very realm (assuming that sets exist). Only when read stipulatively do we have grounds for applying the essentialist inference.

The distinction we have hereby sketched in a preliminary fashion closely resembles the distinction between plain truth and truth by *definition*, or, more generally, between plain truth and conceptual truth. Indeed, one intuitive way of presenting the essentialist inference for sets is to say that since a set has the members it has *by definition*, it *cannot* fail to have just these members. In the next section, we develop these thoughts in more detail.

2.2 Grounding the Essentialist Inference

Up until now, we have been tacitly assuming that the essentialist inference is valid, and concluded that insofar as it is, it crucially depends on the characteristic job of principles of individuation. Hence, in order to decide whether the essentialist inference is indeed any good, we have to understand more clearly what principles of individuation exactly *do*. To that end, we need to take a closer look at the project of metaphysics in general. We already discussed the metaphysical project to quite some extent in the previous chapter, of course. However, there it was our aim to uncover the Modern picture and work towards a sketch of the Aristotelian picture, while here we will consider metaphysics from a different point of view. Here, we are interested in what it is for a metaphysician to endorse one or another list of ontologically fundamental sortal concepts—for that will guide us towards what grounds the essentialist

inference. It is useful to briefly clarify what we have in mind here, by reference to our classification of approaches to metaphysics from chapter 1.

We speak of a 'list' of 'ontologically fundamental sortal concepts' (and below, for brevity, sometimes just of 'fundamental concepts'), and thereby have in mind a list of sortal concepts with a distinctive status to be explicated below. By 'sortal concept', we do not only mean sorts of *things* but any sorts whatever—sorts of properties, sorts of predication, sorts of modality—even sorts of concepts. There may indeed be many sorts of sorts. We thus proceed on a highly generic level, even though we continue to illustrate our claims by reference, largely, to sorts of individual objects—sets in particular. Furthermore, by using the word 'list' we do not mean to imply that it should be an unstructured list. Rather, the list may be structured to any degree. Many metaphysicians speak of 'ontological categories', meaning quite general sorts of sorts—e.g., abstract objects, concrete objects, individuals, tropes, universals, etc. Such ontological categories form the main divisions in the respective metaphysician's list of ontologically fundamental sortal concepts.[92]

Now, endorsing one or another list of ontologically fundamental sortal concepts means something very different for the metaphysical realist than for the metaphysical anti-realist, both of whom work against the background of the Modern picture, and again something different for the metaphysician who operates against the background of the Aristotelian picture. Take, for instance, sets. The metaphysical realist who takes *set* to be an ontologically fundamental sortal concept thereby claims that sets belong to reality as it is in itself (and not just to reality as it is for us). The metaphysical anti-realist, on the other hand, will understand endorsement of an ontologically fundamental sortal concept of sets simply as a matter of convention, as a matter of pragmatic choice, or perhaps as a matter of the inevitable structure of our minds. For the adherent of the Aristotelian picture, endorsing ontologically fundamental sortal concepts comes down to clarifying which concepts are conceptually basic, there being no further question of whether reality-in-itself has a corresponding metaphysical structure or not—the structure of reality *is* that fundamental conceptual structure. Below, we discuss that part of the metaphysical project which consists in providing the allegedly correct list of ontologically fundamental sortal concepts—it is common practice to call that part *ontology*. Throughout, our discussion is meant to

[92] Lowe, for instance, defends a 'four-category ontology', consisting of individual substances, the substantial universals they fall under, the property/relation-instances that characterize them (modes or tropes), and the non-substantial universals these property/relation-instances in turn fall under (which characterize substantial universals) [see Lowe 2006, esp. ch. 2]. By contrast, Campbell [1990], Simons [1994], Keinänen [2011], Keinänen and Hakkarainen [2013], and Giberman [2014] defend a 'one-category' ontology, with individuals conceived of as bundles of tropes and universals as classes of resembling tropes (roughly).

apply to the ontological part of the metaphysical project in all of the three ways just described.

2.2.1 The Ontological Project

If you claim, as a metaphysician, that there are *F*s, you should be able to explain *what* these *F*s are. For suppose someone informs you that her ontology includes *mimsy borogoves*, and that, upon asking what she means, she simply replies: 'I haven't the faintest idea, but they really are amongst the basic furniture of reality'. This obviously doesn't make sense: if we cannot say *what* it is, we should not say *that* it is, and to say what it is is precisely to provide a suitable principle of individuation (*inter alia*). No entity without identity, as Quine famously said.[93]

The ontological task of the metaphysician is to provide and justify an exhaustive list of ontologically fundamental concepts of properly individuated things (or, at least, a sketch thereof). E.g., the nominalist attempts to exclude from his list abstract objects, such as universals and sets; the dualist claims that there are two distinct categories of concrete things, physical and mental; the physicalist claims that everything there is is physical; the Humean will exclude modal properties such as dispositions and powers; the neo-Aristotelian will include a rich variety of basic substantial kinds, probably ordered in some kind of hierarchy, etc. Another part of the metaphysician's job is to show that everything there is ultimately reduces to just these fundamental concepts—which may include denying the existence of some things intuitively held to exist (this is usually called elimination rather than reduction). Thus, the nominalist will argue that properties are not universals but rather particulars (e.g., tropes); the physicalist will argue that the mind is in fact physical (or can be reduced to the physical, or supervenes upon the physical, or doesn't exist); the Humean will argue that dispositions are reducible to categorical properties in some way or other; etc.

Of course, interesting questions can be asked as to the motivations for a metaphysician to suggest one or another ontology, as well as to the credibility and cogency of one or another reductive or eliminative claim. These questions are not our concern in this chapter—they have been to some extent in the previous chapter, and they will be again in Part II of this essay. Rather, we here aim to make a very general point. Take any ontology, and consider any type of things, *F*s, that can be reduced,

[93] See, e.g., Quine [1957, p. 20]. Perhaps reflection on the nature of subatomic particles will reveal that our demand for individuation is too strong—that we should make do, in those cases, with weaker principles of, say, differentiation or relational individuation only (see also our brief remark in fn. 339 on p. 273, and references given there). We do not wish to take a stand on this tricky issue; our argument will apply with only slight modifications also to such a weaker understanding of the ontological project.

according to that ontology—say, the *F*s are just *G*s, e.g., singletons are just sets. Individuation of the *F*s will then be achieved by reference to *G*s—e.g., for singletons x and y it holds good that they are the same just if they are the same set, i.e., have all and only the same members. Let us say that *G*s are *ontologically prior* to *F*s precisely if such a reductive relation obtains. This relation of ontological priority is transitive and asymmetric. We can then take the *ontologically fundamental* sortal concepts to be precisely such that there is nothing they can be reduced to: they are the basic building blocks, metaphysically speaking.

This notion of ontological fundamentality deserves some separate attention. Notice, first, that we assume infinitely descending reductions (without a fundamental level) to be impossible. We shall not argue for that claim here—suffice it to say that virtually all actual ontologies are *not* like that.[94] This assumption, however, does not at all prevent the possibility of a range of fundamental concepts each of which is *accounted for* partly in terms of some (or all) of the others.[95] In such a case, the relevant fundamental concepts will be equidependent on each other. Indeed, we take this to be a very likely scenario, as will become clear in the second part of this essay.

Given an ontology, you have a realm of suitably individuated things at your disposal that can be used for reductions of non-fundamental things along the described lines. When establishing the ontology *itself*, however, no such realm is available. That very realm is thereby being determined, after all. The metaphysician finds herself in the peculiar situation that he is to characterize the ontologically fundamental items and give principles of individuation for them *from scratch*, as it were: the individuation of the ontologically fundamental items is *sui generis*, it does not—indeed *cannot*—depend on the individuation of other, ontologically more prior things.

Earlier, we suggested that there is a factive reading of principles of the form (PI), such as (Set), according to which they simply state something about (e.g.) sets, and that this reading does not provide any grounds for an essentialist inference. A factive reading of (Set) presupposes the existence of sets, which must already come equipped with a proper principle of individuation: it says *of such things* that they are, in fact, identical if and only if coextensive. On this reading, (Set) cannot be used by

[94] One might think that the theory of non-wellfounded sets contradicts this claim, but it does not: both ordinary and non-wellfounded sets are not *reduced* to their members, but only *individuated* by them. What *might* pose a problem for our claim is Ladyman and Ross's contention that 'there might not be a fundamental level', but their notion of fundamentality seems to be a mereological one, and hence quite different from ours. See Ladyman *et al.* [2007, p. 178]. Even 'pandispositionalists' such as Bird [2007] and Marmodoro [2009], whom one may read as claiming that *all there is* is dispositions, such that even the bearers of dispositions are themselves also dispositions whose bearers are again dispositions, and so on *ad infinitum*, do not contradict our claim: their ontology includes just one fundamental concept—dispositions.

[95] It does not even exclude infinite such chains of fundamental concepts.

the metaphysician to present her ontology: it presupposes proper individuation for sets.

So if the metaphysician's ontological project is to make sense at all, there has to be a way for her to present an ontology without using statements read factively. That is, there has to be a *stipulative* reading of principles such as (Set) as well, that amounts, we will say, to taking (Set) as a conceptual truth. (Set), on such a stipulative reading, will be part of the story the ontologist tells us about her fundamental ontological concept of sets. Of course, (Set) will continue to be true on a factive reading as well, but that is not the intended reading of the ontologist. What she tells us about sets is, as it were, true by stipulation, it is *definitive* of sets. Which is just to say that nothing *counts* as a set unless it fits the metaphysician's account, and, *a fortiori*, that nothing counts as *the same* set unless it obeys the principle of individuation included in that account.

It is useful to compare the metaphysician's definition of her ontologically basic concepts with other definitions. We may, for instance, define *bachelor* as 'unmarried adult male human being'.[96] It will then be true by definition, a conceptual truth, that all bachelors are unmarried; nothing counts as a bachelor unless it fits this definition. It is impossible to be a bachelor and yet not to obey the provided definition. Yet bachelors are not essentially unmarried, of course. This is so because bachelors do not *need* to be bachelors: our definition of *bachelor* applies to a realm of ontologically prior, independently individuated things, namely, human beings—their identity does not hinge on their being bachelors, but rather on their being human beings (and if that is not ontologically fundamental, it hinges on their being something else, to which 'human being' reduces). In our terms: bachelors reduce to human beings. Not so for the metaphysician's definition of fundamental concepts: these are supposed to delineate the ontologically irreducible concepts.

We are now in a position to answer certain skeptical worries about principles of individuation, such as those voiced by Michael Jubien:

> There is something very odd about the idea of analyzing identity as it applies to classes (or as it applies to objects of any specific kind). I think it is the same thing that is odd about the idea of analyzing the brother-of relation as it applies to Californians [Jubien 1996, p. 347]
> If identity is a *fully* general, topic-neutral relation, then no matter what kind F happens to be, the correct analysis of "identity as it applies to objects (x and y) of kind F" *must* be of the form '**Id**(x,y) & **F**(x) & **F**(y)', where the first conjunct

[96] We do not wish to claim that this definition coincides perfectly with ordinary usage of the term 'bachelor'.

> analyzes the (general) identity relation *itself*, and the second and third analyze *being of kind F*. [Jubien 1996, p. 348, notationally adapted]

Jubien's objection applies to our account. Identity is as general a relation as one could possibly think of; it applies to anything whatsoever, no matter what kind it instantiates. Hence, Jubien says, it is absurd to think that individuation should be kind-specific—or, as we say, specific to ontologically fundamental sortal concepts. We come to see that being a brother is no different for Californians than it is for other people, so our account of the brother-of relation as it applies to Californians takes the form of a more general account of the brother-of relation that is then restricted to Californians. The thought is, now, that the same applies, *mutatis mutandis*, to principles of individuation, and since, according to Jubien, identity is a primitive notion, identity-for-*F*s must thus be spelled out in terms of that primitive notion plus whatever it takes for something to be an *F*.

Notice that Jubien's understanding of individuation takes all principles of individuation to be reductive: just like the brother-of relation as it applies to Californians is reduced to the more general brother-of relation, so identity for *F*s is reduced to the more general (and primitive) notion of identity. Translated into our terms, this means that Jubien is implicitly endorsing an ontology that incorporates *only one* ontologically fundamental kind of things, best called *bare particulars*, whose individuation is primitive.[97] For any *F* other than 'bare particular', to be *F* just is to be a bare particular satisfying whatever conditions *F*-hood embodies. This result clearly indicates that something has gone wrong—the inference from a seemingly natural observation concerning the generality of identity to such an extreme metaphysical position is just too quick.

What has gone wrong is this: Jubien assumes that the generality of the identity relation is incompatible with there being multiple, informative principles of individuation for diverse ontologically fundamental kinds of things. As his example betrays, he is comparing such principles of individuation to principles of individuation of non-fundamental kinds of things (such as Californians), finds the latter to have the reductive form we described above (for non-fundamental kinds such as 'bachelor'), and generalizes without pause. But compare now, e.g., animals: there are many different kinds of animals, yet they are all animals. Nevertheless, the biologists inform us that what it takes for something to be say, a herring is very different from what it takes for something to be, say, an elephant. There are species-specific accounts of animals, despite the generality of animal. *Indeed, there can be no animal at all without it*

[97] We briefly return to bare particularism in §2.3 below.

being an animal of some specific kind. The same applies to identity: of course everything is self-identical, quite independently of *what* it happens to be—yet for something to be self-identical requires it to be *something*. The ontologically fundamental concepts tell us what forms this 'being something' can, ultimately, take. Jubien's implicit bare particularism is but one of many (*prima facie*) possible answers to the question as to which is the correct ontology.[98]

2.2.2 The Essentialist Inference Vindicated

We may now answer our question: is the essentialist inference applicable to principles of individuation of ontologically fundamental concepts, e.g., of sets? Let us suppose, for *reductio*, that it is not. Now, take a certain set *s*: it could have had different members, according to our supposition. This requires, however, that it makes sense to consider *s* independently from its being the set it is—that is, we are supposing that *s* is individuated independently from its members. Now, there is nothing wrong with that, except for the fact that we thereby are assuming that sets *aren't* ontologically fundamental: apparently, sets are individuated independently from their being (or not being) sets—just as bachelors are individuated independently from their being (or not being) bachelors. That contradicts our assumption that sets make up an ontologically *fundamental* category. So, the essentialist inference applies: if something is a set, and sets are ontologically fundamental, then it cannot fail to be the very set it is.

Every thing is what it is, and not another thing, to use Bishop Butler's famous statement.[99] One may quarrel about which ontology is to be preferred or which one is correct (and we will engage in that quarrel ourselves in Part II of this essay), but however we cut that cake, the essentialist inference is going to apply to the ontologically fundamental concepts—even in the limiting case of bare particularism. That is to say: some form of essentialism is simply inescapable. For the Modern picture-based metaphysical realist, this means that the things that populate reality-itself have essences that somehow correspond to the concepts on the list of ontologically fundamental concepts provided by the (successful) metaphysician; for the anti-realist, this means that the things that populate reality-for-us have essences, which they will be

[98] Similarly, even if van Inwagen [1998] is right in defending the Quinean, 'meta-ontological' thesis that 'the single sense of being or existence is adequately captured by the existential quantifier of formal logic' [Thesis IV, p. 237], this does not mean that what it takes for something to exist can differ amongst different ontologically fundamental concepts. We will defend such a differentiation of existence (into existence, persistence, and living) in Part II of this essay—see, in particular, chapter 4.

[99] See Butler [2009/1827, Preface, §39].

happy to identify with 'our' concepts; for the defender of the Aristotelian picture, this means, quite simply, that essences are concepts that find expression in reality. And that is precisely the tenet of the Aristotelian picture we are defending in this chapter.

To conclude this section, we wish to illustrate the role of principles of individuation by way of an analogy that ties in with the next chapter. Consider model theory: we construct models for theories (or sentences) by choosing appropriate domains and appropriate interpretations for the constants, predicates, and functors in the relevant language. If we don't put any restrictions on these interpretations, the elements of the domain can be understood to be bare particulars—the only thing that holds of them generally is their identity and difference. But if we do put restrictions on the interpretations, and thereby limit the range of admissible models, we get more interesting results. The difference between the stipulative and the factive readings, on which our explication of the essentialist inference rests, is very much like the difference between statements used to delineate admissible models and statements that we interpret within some given range of admissible models.[100]

2.3 Applying the Essentialist Inference

There may seem to be a tension between our highly general claims in this chapter and our focus on the specific example of sets. For sets, being abstract objects, can be clearly and precisely defined and individuated, while the more controversial and interesting applications of the essentialist inference occur not in set theory but rather in more encompassing ontologies. Let us therefore have a look at how our results bear on such ontologies.

An Aristotelian view like the one we will develop and defend in Part II takes the fundamental ontological concepts to include such concrete and specific kinds of things as oak trees, red squirrels, and human beings.[101] Without having to look at the details of such a view, we may already note two things: first, that such concrete objects are indeed very different from abstract ones in that we need empirical investigation in order to get to their essences, and secondly, that we can still have a grasp on part of their essence by way of an Aristotelian story about the genera under which they fall (substance, living being, animal, . . .). The essentialist inference, in this case, can only

[100] We will critically discuss certain aspects of model theory in the next chapter, from which it will become clear what the limits of this analogy are—see especially §3.2.1 and §3.2.2.

[101] Living beings are the paradigm examples of Aristotelian substances, after all.

be properly applied when some significant blanks have been filled in by empirical investigation, but it is nonetheless valid on the same grounds as in the case of sets.[102]

It is more interesting, at this point, to briefly consider another application of our conclusions, to a radically different ontology which we will be critically assessing in much greater detail in chapter 5 below (see, especially, §5.1): that of David Lewis.[103] Lewis has defended, quite brilliantly, a coherent and comprehensive metaphysical picture according to which physical objects are just the contents of arbitrary regions of space-time, such that classical extensional mereology[104] applies to them. In effect, then, he is proposing the following principle of individuation for his fundamental ontological category of physical objects[105]:

(PO) x and y are physical objects $\rightarrow (x = y \leftrightarrow r(x) = r(y))$

(PO') x and y are physical objects $\rightarrow (x = y \leftrightarrow \forall z(O(x,z) \leftrightarrow (O(y,z))))$

($r(x)$ denotes a function that takes objects to the regions of space-time they occupy; $O(x,y)$ denotes the relation that holds between an object and a region of space-time if the former occupies the latter. (PO) states the principle in its intuitive form, (PO') in the form of (PI').)

Being put forward in the ontologist's stipulative mode, this principle licenses the essentialist inference, resulting in the view that objects essentially occupy the region of space-time they in fact occupy:

(PO1) $O(x,r)$

(PO2) $\Box_x(O(x,r))$

(PO3) $\forall x \forall y(O(x,y) \rightarrow \Box_x(O(x,y)))$

This result creates a peculiar tension in Lewis's views. His ontology gives rise, by way of the essentialist inference, to a type of modality, individuation-based modality, that stands in stark contrast with his own notion of counterpart-based modality.

[102] We further clarify the relation between empirical investigation and essential or conceptual truths in the next chapter, in §3.3.3.

[103] See, e.g., Lewis [1986a, 1991].

[104] Classical extensional mereology, or the calculus of individuals, is a formal calculus for the part-whole relation according to which to every number of objects there corresponds precisely one object, their mereological sum, which has all of them as parts and includes nothing that does not overlap any of them. See Simons [1987] for a detailed and critical discussion of classical extensional mereology and various other systems of mereology, and see Mulder [2010] for a discussion of the implications of classical extensional mereology for metaphysics. We revisit the role of classical extensional mereology for certain metaphysical views in §5.1.1.

[105] Quine, whose metaphysical views are quite similar to Lewis's on several scores, has formulated just such a theory of physical objects explicitly *as* his principle of individuation for them; cf. Quine [1985, p. 167] and [1960, p. 171].

Counterpart-based modal truths about some object reduce to non-modal truths concerning that object and its counterparts at other possible worlds in a systematic way (see §5.1.2). The ontological basis for this reduction is given by Lewis's ontology of physical objects (and of possible worlds). Individuation-based modality, on the other hand, arises directly from this ontologically fundamental level. These two conceptions of modality are in conflict: it is counterpart-wise possible that some object occupies a different spatio-temporal region from the one it in fact occupies, yet the very individuation of the objects does not allow for the possibility of such spatio-temporal divergence—different region, different object, after all.[106]

One might consider various ways of dealing with this tension, on behalf of Lewis. For example, one might point out that individuation-based modality is 'intra-world' while counterpart-based modality concerns 'inter-world' relations (of similarity), and that therefore there is no real tension between the two notions. We would then argue that the very existence of the mentioned individuation-based modality within Lewis's own system shows that his attempted reduction of modality fails, and that the whole counterpart-theoretic story is in fact a distraction that obscures the real nature of Lewis's physical objects.[107] Anyway, attempts to absorb individuation-based modality into the counterpart-theoretic framework are bound to fail because the counterpart-theoretic framework rests upon an ontologically fundamental foundation which comes with its own individuation-based modality. For our purposes in this chapter it is enough to simply indicate the far from trivial consequences that our proposed understanding of the essentialist inference have, by way of some interesting examples.

Another such example is the view known as *bare particularism*, which we mentioned before in discussing Jubien's skepticism about principles of individuation (see §2.2.1, p. 73 above). We choose this example, despite all the delicate problems that it invites, because it illustrates the limiting case of the essentialist inference. On at least some versions of the doctrine (in particular, the one we ascribed to Jubien), every existing thing is, at bottom, a bare particular which exemplifies a range of

[106] A Lewisian might argue that he has a region-of-space-time counterpart relation at his disposal, which picks out for a given object only counterparts that occupy regions of space-time exactly similar to the one it in fact occupies, and thus yields exactly the same statements that we have given—(PO2) and (PO3). However, this would be a rather superficial remark: the content of the counterpart-based versions of (PO2) and (PO3) would be different from the content they have on an individuation-based account. Different modal notions are involved. Syntactically, the difference shows up with the small suffix x we have attached to the $\Box$.

[107] Heller [1990] bites this bullet: he defends a Lewis-like metaphysics which explicitly includes commitment to spatiotemporal (and mereological) essentialism. For a more thorough discussion of the issue, see Mulder [2010], and our discussion of Humean ('first-gear') metaphysics in chapter 5.

properties none of which is essential to it. The view in effect proposes an ontological reduction (in our sense) of everything to bare particulars: for existing things across the board, it holds that $x = y$ just if they are (or 'have') the same bare particular. Identity is *always* spelled out in terms of identity between bare particulars, whose identity and difference is, in turn, a brute fact.[108] It is important to note that this does not result in an ontology *without* essentialist commitments—it just makes these commitments very shallow and uninteresting. Take any random object, say, Socrates. He *can* be (identical to) anything (a set, a maple tree, a galaxy, . . .), but he will always be the self-same bare particular, in accord with a principle of individuation like the following:

(BP) x and y exist $\rightarrow (x = y \leftrightarrow b(x) = b(y))$

(BP') x and y exist $\rightarrow (x = y \leftrightarrow \forall z(z = b(x) \leftrightarrow z = b(y)))$

($b(x)$ denotes the bare particular of x; (BP) states the principle in its intuitive form, and (BP') in accordance with (PI')). Along familiar lines, we may conclude on the basis of (BP') that any existing thing has its bare particular essentially—and nothing more.[109]

The bare particularist has to say that what it is to be (say) a set, is just to be a bare particular satisfying certain conditions that it could also fail to satisfy (say, being an abstract object to which the membership relation applies), just as most ontologists will say that what it is to be a bachelor (as defined earlier) just is to be a human being satisfying certain conditions (i.e., being unmarried, adult, and male).

For the not so austere ontologist, who defends a more interesting list of ontologically fundamental concepts, it is impossible to deny the essentialist inference for the concepts that he includes in his list without implicitly changing his list and thereby moving towards bare particularism. And even the bare particularist himself cannot avoid a residue of essentialism, however trivial that may seem to be (i.e., essential self-identity and other-distinctness).[110] Thus we conclude that essentialism is not an

[108] For recent defenses of bare particularism as a metaphysical view, see Moreland [1998, 2001], Mackie [2006], Sider [2006a], and Davis and D. S. Brown [2007]—but, to be sure, none of them will accept our toy version of the doctrine. For recent critical discussion, see Mertz [2001, 2003] and Bailey [2010].

[109] As we said, the discussions on these matters are fraught with difficulties. It is argued, for example, that Socrates is still essentially human, on a certain understanding of bare particularism, because Socrates needs to be human even though his bare particular need not. Or even that his bare particular necessarily comes with the property of being human. See, e.g., Moreland [1998, 2001] and Pickavance [2009]. We confess our inability to see how these claims can be coherently endorsed within a bare particularist framework—they seem to amount to abandoning bare particularism.

[110] For a powerful defense of the same claim from a rather different point of view, see Fine's detailed analysis of Quine's arguments against 'quantifying in': Fine [1989, 1990, 2005c].

optional ingredient in one's metaphysical views: it is constitutive of the very project of metaphysics.

2.4 Essences and Concepts

As we have seen, there is no way to pull any essentialist conclusions out of the hat of the mere statement of a principle of the form (PI) or (PI'). By reflection on the job of the metaphysician, however, we have found out that what he says about the ontologically fundamental concepts is, in an important sense, different from everything we may then go on to say or ask about instances of these concepts. While presenting his list of fundamental sortal concepts, the metaphysician goes into the *stipulative mode*, so to speak; once they have been laid out, we can go into the *factive mode* and consider them and their instances.

What the metaphysician tells us about the entities there are provides the background conditions for engaging in discussions about these entities, for considering what is, can or must be true of them. So if the metaphysician says that sets are individuated by their members in accord with (Set), this is meant to be part of these background conditions and hence not a matter of dispute once we have accepted the ontology. We may of course question the ontology, but that's an entirely different game.

We have not elaborated on our sketchy formal reconstruction of the essentialist inference. Although this can surely be done (cf. the proposals in this direction by van Cleve [1985]), it would not add much: the principles one needs to invoke in order to lay down such a formal reconstruction can only be justified by reference to considerations such as the ones we have given concerning the metaphysician's ontological project.

In our discussion of the job of the metaphysician we have largely abstained from comparing it with our conception of metaphysics laid out in the previous chapter; we merely noted that our discussion would apply both to those working within the Modern picture, be it as realists or as anti-realist, and to those working within the Aristotelian picture. Let us now abandon this neutral stance and have a look at the relation between concepts and essences that our understanding of the project of ontology implies.

In providing a list of ontologically fundamental concepts and their interrelations, the metaphysician in effect lays down a specific set of sortal concepts that differ from

other sortal concepts in that they are ontologically fundamental.[111] Consider, again, the difference between the sortal concepts *set* and *bachelor*. Assuming the former but not the latter to be a ontologically fundamental concept, we may observe that if we say truly of something that it is a set, we have thereby arrived at the most fundamental understanding of what it is we have before us, while if we say truly of something that it is a bachelor, we have not arrived at such a point. This is because the concept *bachelor* presupposes a realm of things to which it can be applied, namely, human beings, while the concept *set* does not, as we have seen.[112]

By way of illustration, let us consider a kind of statement that Aristotle favored: *qua*-statements. *Qua* bachelor, this man is necessarily unmarried, one might say, while *qua* lawyer, he is not necessarily unmarried. This divergence in modal predication constituted a major motivation for Quine to endorse his modal skepticism, and for Lewis to endorse his counterpart-theoretic account of modality.[113] We can now see that such divergence in modal predication is unproblematic because the man has his identity independently from his being a bachelor (and from his being a lawyer); there is a more fundamental answer to the Aristotelian question *What is it?* than 'bachelor'—let's say, the fundamental answer is 'human being'. Now, we may also say that, *qua* human being, this man is necessarily mortal. Since the man's identity is *not* independent from his being human—*human being* by assumption being an ontologically fundamental concept—there is no alternative way of regarding the man which will make it false that he is necessarily mortal. There is a distinction, then, between what one might call *hypothetical qua-statements*, where the sortal concept *qua* which we are considering something is a non-fundamental one, and what one might call *categorical qua-statements*, where the sortal concept *qua* which we are considering something is a fundamental one. This distinction ties in with the distinction between *de re* and *de dicto* modality: if you replace the subject of a hypothetical *qua*-statement with its *qua*-part, you get a statement which is only true *de dicto*, while if you replace the subject of a categorical *qua*-statement with its *qua*-part, you get a statement which is true *de re* as well. E.g., '*qua* bachelor, this man is necessarily unmarried' becomes 'this bachelor is necessarily unmarried', which is false *de re* but true *de*

[111] Alternatively, one might say that only the ontologically fundamental sortals are 'real' sortals—but that is a purely terminological issue. For discussion of the distinction between fundamental and non-fundamental sortals, see §4.2, p. 133.

[112] To be sure, one might also argue that by saying of something that it is a bachelor, we *have* arrived at the most fundamental understanding of what it is: being a bachelor entails being a human being, after all, and that is going to be (or guide us towards) the fundamental concept under which the given thing falls. Again, see §4.2, p. 133 for further discussion.

[113] See, e.g., Quine [1960, pp. 199–200] and Lewis [1986a, pp. 248–63].

dicto; '*qua* human being, this man is necessarily mortal' becomes 'this human being is necessarily mortal', which is true *de re*.

Sortal concepts imply a range of conceptual truths—truths like *bachelors are unmarried* and *human beings are mortal*. Such truths carry a modal force which derives from the sortal concepts involved; the modal force is the same in both of our examples. Yet there is a difference between them, which, on the Aristotelian picture, comes down to the following: some sortal concepts are essences, others are not. It is of central importance to our ongoing search for a better understanding of (parts of) reality, be it in everyday life or in scientific investigation, to find out which are the ontologically fundamental sortals. It is the metaphysician's job to reflect on this question on a very general, conceptual level. That is what we aim to contribute to in Part II.

To repeat: on the Aristotelian picture, fundamental sortal concepts simply *are* essences—general essences, to be more precise. If we grasp them correctly, what we grasp is identical with what is 'out there'. Grasping the concept *set* means grasping the essence of sets in general; grasping the identity of some specific set means grasping the individual essence of that particular set. As we promised in §2.1.1 (see fn. 82 on p. 63), we can now say a bit more about our understanding of such individual essences.

On the Aristotelian picture, the relation between individual things and their individual essences is identity. That is to say, an individual essence is not *one among the elements constituting* an individual thing, but it rather *is* that thing. Put very briefly, this identification is required because otherwise the work essences are supposed to do, namely to unify that of which they are essences, is always incomplete: the essence is supposed to unify the composing elements, of which it is just one, but it is thereby left open how these elements themselves are unified: what are *their* essences? A regress threatens.[114] To understand what that means, consider the explanation Anna Marmodoro provides:

> The step ... from the level of the abstract essence 'human being' to the level of the individual human being has two aspects to it. First, it is a difference of a determination level. Second, it results in the maximal determination, namely in the level of the *determinate* rather than of a further determinable. [Marmodoro 2009, p. 37]

[114] Arguments of this style go back to Aristotle himself, who argued in favor of the identification of entity with individual essence in *Metaphysics* VII. There, he concludes that 'with things that are primary and spoken of *per se*, the [essence of the thing] and the thing itself are one and the same' [Aristotle 1998, p. 188, 1032a2–4; we replaced the awkward 'what-it-was-to-be-that-thing' with 'essence of the thing']. See Scaltsas [1994] for an elaborate interpretation of Aristotle that is sympathetic to our understanding.

The thought is, roughly, as follows. Starting with the concept (general essence) of a human being, one may notice that it includes a host of determinables that one may specify to obtain more specific concepts—e.g., dark-haired human being, seated human being, happy human being. Such more specific concepts are reducible, of course, to the concept of a human being itself. Now, one may play this specification game indefinitely long, refining and specifying the resulting concepts with more qualifications or combinations thereof. However, as soon as one adds a specification that constitutes (or implies) individuation for human beings, we have moved from the general realm to the realm of the particular: we are no longer thinking about *a* human being satisfying certain conditions, but rather about *this particular* human being. Once we have arrived there, we find that nothing is indeterminate anymore: our particular human being has a specific hair color, a specific posture, is in a certain specific mood, etc. That is the difference in 'determination level' typical of individual essences: maximal determination.[115]

One might go on, here, to add remarks on the Aristotelian notion of matter, which 'marks the level of full determination of the essence primitively, without giving matter the role of a resident particular in the constitution of the thing' [Marmodoro 2009, p. 37]. That, however, would take us too far afield, and, moreover, would seem to focus our attention too much on concrete objects and living organisms. Instead, to stress the generality of the proposed understanding of individual essences, notice that things are no different for abstract objects. Take the concept (general essence) of natural numbers: one may add specifications to yield the concept of an odd number, of an odd prime number, of an odd twin prime, of an odd twin prime which succeeds 126—and, of course, this last step implies (or, rather, constitutes) individuation of the (yes, now *the*) natural number: 127. Determinateness in all respects is the mark both of the particular object and of the particular essence—which is no surprise, given that they are one and the same.

In the next chapter, we turn to the concepts-side of our central thesis that essences are concepts. Here, we have developed an understanding of essentialism that rests on the identification of essences with ontologically fundamental concepts. Our mission for this chapter has thus been accomplished: we may conclude that essentialism is true—and not only on the Aristotelian picture, we should add, but on any construal

[115] That the *individual* essence is maximally determinate does not imply that all these determinations are essential to it—e.g., this particular human being has brown hair, but that is not essential to him. The relevant essential truths are those that follow from the *general* essence it instantiates. These may include certain determinate truths, but they can leave open many determinations as well—thereby leaving room for contingent matters of fact.

of the metaphysical project. Essentialism is a precondition for metaphysics as such. In the next chapter, we zoom in on the question how concepts guarantee truths.

* * *

Chapter 3

Conceptual Truths

It is our aim in this chapter to further defend and elucidate our claim that essences are concepts. Whereas the previous chapter did so by focusing on the essences-side of that claim, so to speak, this chapter is devoted to an investigation into the concepts-side. We do so by investigating the phenomenon of conceptual truth. Conceptual truths express how concepts hang together, hence they also underlie the very idea of something's following from something, that is, the very idea of valid inference. We claim that this latter notion relates to conceptual truth in precisely the way in which logical consequence relates to logical truth, and, moreover, that these more specific, logical notions in fact result from a restriction on conceptual truth and consequence.

A lot of work has been done, especially over the past century, on the topic of logical consequence and logical truth, with the hope, one may assume, of arriving at a well-understood, precise construal of valid inference. The present chapter starts with a focused investigation of one popular approach to these logical notions in the contemporary debate: the Tarskian approach. The lessons we then draw from this investigation complement those of the previous chapter such as to yield a substantive elaboration of the metametaphysical reading of our claim that essences are concepts, which forms part of the foundation of the Aristotelian picture as we outlined it in Chapter 1.

Traditionally, the notions of logical consequence and logical truth are taken to be purely *formal* notions: logic concerns logical form, in abstraction from content. What we call the Tarskian approach to logical consequence and truth takes this

formality as its point of departure and proceeds to offer reductive definitions of these logical notions, in terms of truth under arbitrary reinterpretation of all non-logical expressions. What remains, on this view, is the question as to the proper distinction between the logical and the non-logical terms, or, in other words, between form and content: that distinction demarcates logic and hence also the notions of logical truth and consequence. We, however, argue that the search for such a principled demarcation is misguided. The error, we claim, lies in the confused idea of form *as opposed to* content. Understanding formality in this Tarskian manner also threatens to obscure the modal element that characterizes the very idea of valid inference, and hence also its specifically logical instance, the relation of logical consequence. This modal element, however, turns out to be tacitly presupposed even by the Tarskian view itself.

We conclude that *logical* truth and consequence indeed are restrictions on a pair of more basic notions, viz., *conceptual* truth and consequence. These latter notions are themselves best spelled out *not* in the Tarskian way, but rather in terms of *truth being guaranteed by content alone*. Although closely tied to the traditional notion of analyticity, our understanding of conceptual truth as truth grounded in content is different, which is not surprising given that we understand concepts from the point of view of the Aristotelian picture (as developed in §1.3). We develop in rough outline a conception of form that treats it as a restriction on the allowed content: form is what you get when you leave out *certain* kinds of content—not *all* content. The traditional notion of logical form can then be understood, roughly, to be the form which results if you leave out all content but that contributed by the logical constants (whichever these are taken to be). Such a construal of logical form gives rise to notions of logical consequence and truth as restrictions on the broader notions of conceptual consequence and truth. The study of such restricted notions, though very worthwhile as such, cannot provide materials for reductively defining the very idea of conceptual truth, or of something's following from something (as the Tarskian approach to logical consequence seemed to promise), since such restricted notions are themselves still to be understood in terms of truth guaranteed by content.

Truths whose truth is guaranteed by their logical form are logical truths. They can be said to be true by *logical necessity*, just as logical consequence can be said to be a relation that holds, if it does, by logical necessity. In the same vein, we can say that *conceptual* truths, whose truth is guaranteed by their content without any restriction, are true by *conceptual necessity*. Unsurprisingly, the notion of logical necessity turns out to be a restriction on the more fundamental notion of conceptual necessity.

Having thus clarified the interrelations between the concepts we are interested in, we notice that our central notion of conceptual truth, of truth grounded in content, strikingly resembles the notion of essential truth as grounded in the identity of things, which we borrowed from Fine and developed in our own way in the previous chapter. Hence logical and conceptual investigation, on our understanding, simply *is* metaphysical investigation. Thus we complete our outline of the metametaphysical foundation which underlies the second part of this essay: there, we will have occasion to embark on precisely such logical/conceptual investigations, though the types of content we will be interested in there outstrip those falling under the heading of logic as traditionally understood.

We proceed as follows. After a brief introduction (§3.1), we consider John Etchemendy's well-known attack on the Tarskian approach to understanding logical consequence (§3.2). While agreeing with Etchemendy on the shortcomings of this reductive approach, we conclude that two fundamental questions remain to be answered: what sense can be made of the idea that logic is 'formal', and what exactly is the modality involved in the consequence relation? These questions will occupy us at length in §3.3, where we approach them with the help of a number of ideas we borrow from Kit Fine—some of which we have already acquainted ourselves with in the previous chapter. There, we develop our notion of truth guaranteed by content, which gives rise to the notions of conceptual truth and conceptual consequence, in terms of which logical truth and consequence can then be understood. On this basis, we argue that, from a metaphysical point of view, logical orthodoxy has created an unfortunate bias when it comes to the kinds of content it focuses on (§3.3.2): moving beyond that logical orthodoxy, we pave the way towards the conceptual investigations which we will engage in in the second part of this essay. Along the way, we find occasion to reflect on how our understanding of conceptual truth, conceptual consequence, and conceptual modality relates to the traditional notions of analyticity and aprioricity (§3.3.3). We conclude with a brief summary in §3.4.

3.1 Valid Inference and Logical Consequence

The consequence relation is, in one good sense, *the* core concept of philosophical logic—that is, insofar as we conceive of philosophical logic as the study of correct reasoning, of what follows from what, or as the science of thought. The consequence relation holds between some given premises and a given conclusion whenever the conclusion *follows from* the premises, whenever the premises *entail* the conclusion,

whenever the truth of the premises *guarantees* the truth of the conclusion. A given argument or inference is valid if the consequence relation holds between its premises and its conclusion. The consequence relation, then, should guide our thinking. Grasping the conclusions one may validly draw from the thoughts one is entertaining amounts to a better understanding of those thoughts themselves.

Now, since we approach the very questions of logic with our capacity for thought (with what else?), it is fair to say that one central task of philosophical logic is to elucidate or explicate the relation of consequence which we rely on in the very process of thinking about it, while not abstaining from revising that notion where inconsistency threatens (and yes, consistency itself is, of course, a close cousin of the notion of validity).

That said, let us now turn to some preliminary remarks on the consequence relation as it is captured in (contemporary) logic. We quickly learn that we are thereby restricting our attention to a more specific notion, to be defined within the artificial languages of various logics. This notion of *logical consequence* can be very precisely defined, has been extensively studied, and is thus much better understood than the intuitive consequence relation. It is hoped, then, that our sophisticated understanding of logical consequence can help us in understanding the intuitive consequence relation—or even, that for serious scientific or philosophical purposes, this well-defined notion of logical consequence is all we need.

What we need, first, is a way to be explicit about the difference between consequence in general and logical consequence in particular. Let us say that the intuitive consequence relation holds between thoughts or contentful sentences: we denote such entities with capital letters A, B, C etc., and classes thereof with calligraphic capitals $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ etc. We will express the intuitive consequence relation with the symbol $\triangleright$, so that we can state that a conclusion A follows from a class of premises $\mathcal{A}$ by writing $\mathcal{A} \triangleright A$. By contrast, the relation of logical consequence holds between formulae of a well-defined logical language, which are intended to capture the logical form (on one or another construal thereof) of contentful sentences. For such formulae, we use lower-case Greek letters φ, ψ, etc., and for classes thereof we use upper-case Greek letters Γ, Δ etc. To explicate the relations between contentful sentences and formulae, we write $\varphi = \overline{A}$, to state that φ is the logical form (on some construal thereof) of A, or $\Gamma = \overline{\mathcal{A}}$ to state that Γ contains all the logical forms (on some construal thereof) of the members of $\mathcal{A}$, and no others.[116]

[116] We leave open whether or not contentful sentences admit of multiple readings. If so, there may be multiple logical forms (on some construal thereof) that correspond to one and the same contentful sentence.

We may still say of a contentful sentence A that it follows *logically* from a given set of premises $\mathcal{A}$, on the basis of one or another well-defined logical language. By doing so, we express that the relation of logical consequence holds between Γ and φ in that language, where $\Gamma = \overline{\mathcal{A}}$ and $\varphi = \overline{A}$.

Basically, an argument couched in terms of a given formal language consists of a set Γ of formulae in that language, the premises, and another formula φ, the conclusion. The relation of logical consequence holds, between premises and conclusion, if and only if the argument is *truth-preserving* in a formal way. Put loosely: if the premises are all true, the conclusion must be true as well—and this should hold for all variations on the premises and conclusion that preserve their logical form. This, then, is the core of the notion of logical consequence. The related notion of logical truth can be defined to be a property which a formula φ has if and only if it follows from no premises at all, that is, if and only if its truth, and the truth of all sentences of the same form, is unconditionally guaranteed.[117]

As we teach our first-year students, there are two ways of spelling out this notion of logical consequence. On the one hand, there is the syntactical approach which centers around the notion of proof or derivation, for which several systems of rules and/or axioms exist. The intuitive idea is that one aims to show, by using securedly valid axioms and rules of inference, how to arrive at the conclusion on the basis of the premises. This syntactic rendering of logical consequence might be called *derivability* or *provability*. On the other hand, there is the semantical approach which centers around the notions of a model and of truth-in-a-model. The intuitive idea is that one aims to show that the argument in question is truth-preserving by showing that in all situations in which the premises are true, the conclusion is true as well. This semantic rendering of logical consequence might be called *semantic validity*. The syntactic conception of logical consequence, provability, is denoted with $\vdash$, the semantic conception with $\models$. And then we have, at least in the case of some logics, soundness and completeness proofs which show that the two resulting conceptions of logical consequence do not diverge for those logics.[118]

In that case, our function-like form-operator ($\overline{A}$) is inappropriate. We ignore this possible complication for simplicity's sake; what we have to say does not depend on this simplifying assumption. Also, we omit explicating subtleties concerning mention and use: when writing $\varphi = \overline{A}$ we are mentioning both φ and A, not using them.

[117] To be more precise: 'having the same logical form' is normally taken to consist in containing exactly the same occurrences of logical constants and the same pattern of occurrences of non-logical terms.

[118] Terminological issues are difficult. Some use 'logical consequence' to mean what we call semantic validity here, that is, to denote the semantical conception of logical consequence; some use 'validity' to denote that same semantical notion. Just to be explicit: we use 'the consequence relation' or 'validity' to denote the very idea of something following from something ($\rhd$); 'logical consequence' or 'logical

When introducing logical consequence, or indeed the consequence relation as such, one usually attempts to get it across both in modal and in epistemic terms. Thus, it is said that if A (logically) follows from $\mathcal{A}$, it is *impossible* that all of $\mathcal{A}$ be true while A is false, or that if all of $\mathcal{A}$ are true, then A *must* be true as well.[119] We will say that a valid argument *guarantees* the truth of the conclusion provided the premises are all true—and, indeed, we will be concerned with the nature of such guarantee. On the other hand, it is said that if A (logically) follows from $\mathcal{A}$ and we know all of $\mathcal{A}$, we may safely conclude that A.[120] We prefer to explain the epistemic justification that is thus claimed for this inferential step in terms of the modal characterization: if the premises are true, then the conclusion is *guaranteed* to be true as well.

So much for preliminaries. We are interested in logical consequence because we hope to arrive at a better understanding of the consequence relation in general—by developing a suitable understanding of that relation, we aim to arrive at an understanding of conceptual truth that sheds light upon our central claim: that essences are concepts.

3.2 A Reductive Approach

One firm tradition in the philosophy of logic, pioneered amongst others by Alfred Tarski, holds that logical consequence and logical truth are best understood *reductively*, in model-theoretic terms. We will call this the *Tarskian approach*: it consists, at root, in a reduction of logical consequence and logical truth to ordinary, 'material' truth. The idea is, then, to take the semantical construal of logical consequence as fundamental. We are all familiar with this idea: φ is logically true (or 'universally valid') just when it is true in all suitable models, and φ is a logical consequence of Γ just when it is true in all suitable models in which all of Γ are also true.[121]

The Tarskian approach is deeply ingrained in our thinking about logic. One symptom of this, as John Etchemendy observes[122], is our labeling the soundness and

validity' to use the specifically logical instance of that idea, 'provability' to denote the syntactical construal of logical consequence in terms of the existence of a proof ($\vdash$), and 'semantic validity' to denote the semantical construal in terms of a model-theoretically defined relation of truth-preservation between premises and conclusion ($\models$).

[119] Etchemendy provides a number of quotes from introductory, intermediate, and advanced texts, to illustrate the ubiquitousness of this modal characterization of the consequence relation. See Etchemendy [1990, p. 82].

[120] See, for instance, Quine's remark on logical consequence in his introductory text on logic [Quine 1950, §7].

[121] A somewhat more detailed exposition of this Tarskian approach will be provided shortly (in §3.2.1).

[122] See, e.g., Etchemendy [1983, p. 325, 1988, p. 67, 1990, p. 3].

completeness theorems as we do:

$$\textsc{Soundness} \quad \text{If } \Gamma \vdash \varphi \text{ then } \Gamma \models \varphi \qquad \textsc{Completeness} \quad \text{If } \Gamma \models \varphi \text{ then } \Gamma \vdash \varphi$$

We use the soundness theorem to show that some form of provability, $\vdash$, which is a syntactical construal of logical consequence for a given logic, yields *only* valid inferences, while we use the completeness theorem to show that $\vdash$ captures *all* valid inferences. Apparently, then, we treat $\models$ as the standard against which we evaluate $\vdash$. It is not at all inconceivable, however, that these theorems are used the other way around, i.e., by using a given notion of provability, $\vdash$, which is taken to faithfully capture all and only valid inferences in the given language, to test $\models$ against. One would then use *Soundness* to show that $\models$ yields all relevant valid inferences (i.e., that $\models$ is *complete*), and *Completeness* to show that $\models$ yields only valid inferences (i.e., that $\models$ is *sound*). The fact that we are so used to the former way of reading these theorems betrays how deep our thinking about logical consequence has been shaped by the Tarskian approach: we assume that model theory has the last word when it comes to the extension of the relation of logical consequence.[123]

The strategy that is employed in the Tarskian approach is one of reduction by generalization, as we may call it. It is a very familiar strategy, which we also find in other areas of philosophy—for example, in certain broadly Humean accounts of causation, which aim to reduce the truth of particular *causal* statements to the truth of corresponding generalized *non-causal* statements: 'this *a* caused that *b*' is true, on such accounts, because its non-causal counterpart 'this *a* is followed by that *b*' instantiates the exceptionless generalization '*a*'s' are always followed by *b*'s' (roughly). Here as in our case, what disappears from the scene once the reduction is in place is the modal aspect of the original notion: it is replaced by a (meta-level) generalization. The Humean thus purportedly gets rid of causal modality; likewise, the Tarskian purportedly gets rid of what we call logical modality. And in both cases, there is reason to resist such reduction precisely because of the modality involved.[124]

This is where Etchemendy's [1983, 1988, 1990] forceful campaign against the Tarskian approach comes in. Etchemendy provides both an in-depth analysis of

[123] Georg Kreisel was very much aware of this situation; in fact, he provided an (informal yet rigorous) proof that, for first-order sentences φ, the intuitive notion of consequence $\triangleright$ extensionally coincides both with the model-theoretic notion of validity $\models$ and with the syntactical notion of derivability $\vdash$. See Kreisel [1967, p. 154f].

[124] Other examples of such a strategy include David Lewis's reductive account of modality (which we encountered already in §2.3), and broadly Humean accounts of laws of nature. In the second part of this essay we will have occasion to discuss such reductive projects in much more detail—see, for instance, 5.2 for a discussion of Humean accounts of laws, and 6.1 for a discussion of causation. See also Stroud [2011] for a forceful resistance to such reductive projects.

Tarski's original model-theoretic definition of logical validity (see Tarski [1936]) as well as a critical discussion of the refined version that is current orthodoxy. We will draw some important lessons from Etchemendy's work, for which we need to dive a little bit deeper into the Tarskian conception of logical consequence (both into Tarski's original proposal and into the contemporary version).

Etchemendy makes two interrelated points which we will explore below. The first point is that the Tarskian approach fails to be extensionally adequate, i.e., declares some arguments valid which in fact are not valid, or fails to declare some arguments valid which in fact are valid. Etchemendy calls these problems of 'overgeneration' and 'undergeneration', respectively. The second point, which relies heavily on the first, is that the model-theoretic conception of logical consequence as semantic validity, $\models$, gets the order of explanation wrong: the truth of the matter is that we construct model theory the way we do because we want it to capture our *presupposed* notion of consequence as it applies to arguments couched in the language at hand. That is, if φ follows logically from Γ, we should view the fact that there are no models which make φ false but all of Γ true as evidence that our model theory is adequate, and not as grounding the very fact that φ follows from Γ. In short: reducing logical consequence to $\models$ is very much like putting the cart before the horse.[125] Which is not to say that we should instead go for some kind of provability, that we should adopt a purely syntactic conception of logical consequence instead of a Tarskian, semantic conception. We will come back to this issue in §3.3 below.

3.2.1 Extensional Adequacy

Tarski's point of departure is a set of two conditions on any definition of the notion of logical consequence: it should *preserve truth* and it should be a *formal* relation. Tarski initially glosses this second requirement as follows: if A follows logically from $\mathcal{A}$, then uniformly replacing all the non-logical constants in all of $\mathcal{A}$ and in A should not affect the consequence relation.[126]

However, truth-preservation and formality in this sense are together not sufficient for capturing logical consequence, because the language for which we are considering it might simply not contain enough expressions. For example, take a language containing only one singular term, 'Rembrandt', and one predicate expression, 'x was a painter'. Within this language, the sentence 'Rembrandt was a painter' would

[125] Again, there is an interesting parallel to be drawn to broadly Humean, reductive approaches to laws of nature and to causation—as we point out in §5.2.2, p. 171 and §6.1, p. 184, respectively.
[126] See Tarski [1936, pp. 414–5].

then be a logical truth, because it is in fact true and remains true under all substitutions of non-logical terms—only the identity substitution is available, after all.[127]

The solution to this rather technical problem is found in the semantical notion of *satisfaction*. Instead of replacing non-logical constants with other constants from the language, we can also replace them by (meta-linguistic) *variables* of appropriate types. Thus we obtain from the class of sentences $\mathcal{A}$ the class of sentential functions $\mathcal{A}^*$. Tarski now defines a sequence s of objects that satisfies all these sentential functions to be a *model* of $\mathcal{A}$. Now the concept of logical consequence is defined as follows:

> The sentence [A] *follows logically* from the sentences of the class [$\mathcal{A}$] if and only if every model of the class [$\mathcal{A}$] is also a model of the sentence [A] [Tarski 1936, p. 417; notationally adapted].

Truth-preservation is directly built into this definition, and the requirement of formality is now dealt with quite elegantly: Tarski circumvents the possible expressive weakness of a logical language by moving from a syntactic strategy—substitution of non-logical constants—to a semantic strategy: reinterpretation of non-logical constants. However, Tarski ends his discussion of these matters by noting one potential weakness of his definition: it is relative to a set $\mathfrak{F}$ of terms, the logical constants, which are not to be reinterpreted.[128]

Without taking a stance on the question whether or not Tarski meant his semantical definition of logical consequence to be a reductive definition, we may say that it allows for a reductive reading. It is easiest to see how this reduction works if we focus on Tarski's related notion of logical truth, which is defined as truth under arbitrary reinterpretation of all the non-logical constants. First, we abstract away from the actual contents of a candidate logical truth by focusing solely on its *logical form*: the occurrences of logical constants from $\mathfrak{F}$, plus the pattern of occurrences of the remaining expressions. Consider, for example, the following sentence:

If Rembrandt was a painter, then Rembrandt was a painter.
$Pr \rightarrow Pr$

Then, assuming $\rightarrow$ to be included in $\mathfrak{F}$, we find out that this sentence is logically true, because it is true under arbitrary reinterpretation of P and r—no matter how we assign extensions to P and r, the resulting sentence is true. Thus, our sentence's being logically true comes down to the following (higher-order) generalization's being *materially* true:

[127] See Tarski [1936, pp. 415–6]. Notice that Tarski is working with interpreted languages here, which is required for the substitutional approach to make sense—hence our use of A and $\mathcal{A}$ rather than φ and Γ.
[128] See Tarski [1936, pp. 418ff.].

$\forall\Phi\forall\alpha(\Phi\alpha \rightarrow \Phi\alpha)$

(Where Φ is a metalinguistic variable for properties and α is a metalinguistic variable for objects.) To repeat: this reduction of logical to material truth is relative to a set $\mathfrak{F}$ of logical constants that are *not* to be replaced in the described way.

In our terms, we may say that Tarski here moved from contentful sentences A, B etc. and substitutional variations thereof to something much more like uninterpreted logical formulae γ, ψ etc.

So far, we have just been looking at Tarski's original definition of logical consequence and logical truth. On this definition, the domain from which the sequences of objects are taken is not explicitly specified, hence it is implicitly taken to simply contain all and only the actual objects. Thus, it is invariant. Contemporary model theory has a more elaborate notion of a model which includes a variable domain of quantification. And with good reason: the Tarskian definition, which tacitly takes all the actually existing objects as its domain, results in sentences like $\exists x\exists y\exists z(x \neq y \wedge y \neq z \wedge x \neq z)$, saying that there are at least 3 objects, being declared logically true (provided that we include = in $\mathfrak{F}$). Such sentences do not contain any non-logical constants, hence if they are true, it is quite trivial that they are also true under reinterpretation of all non-logical constants. Varying the domain of quantification eliminates such overgeneration of logical truths.[129]

Etchemendy now complains that this refinement seems a very *ad hoc* solution, and that the resulting construal of logical consequence leaves important questions unanswered.[130] For example, are we still treating $\exists$ as included in $\mathfrak{F}$? Don't we in this way simply reinterpret $\exists$ by assigning other (more or less restricted) quantifiers to it, hence treating it as a non-logical constant (i.e., excluding it from $\mathfrak{F}$)? Where Tarski's original account reduces logical truth to material truth, that is, to truth in the actual world, the present modification seems to reduce it to truth-in-all-'subworlds', that is, to truth within arbitrarily chosen subclasses of the class of all actually existing things. And, even if such worries can be set aside, there are still problems with extensional adequacy. For consider finitism, the metaphysical view that there are only finitely many objects—say, n. On this view, the domain of quantification has a maximal cardinality of n, hence a sentence saying that there are no more than n objects will

[129] Modern model theory also comes with a different view on the logical language: where Tarski was operating with interpreted languages, such that a detour through sentential functions involving variables of diverse types is required for his satisfaction-based method to work, modern model theory normally operates with a purely *formal* understanding of logical languages, such that the uninterpreted predicate symbols and individual constants themselves play the role of these variables. That is where we move from contentful sentences (A) to uninterpreted formulae (φ).

[130] See Etchemendy [1990, esp. chs. 5 and 8].

be judged logically true: no actuality-based countermodel can be given in which it is false.

At this point, one might think that we are being unfair in insisting on models' being populated by *actual* objects. Why not take the models to be possible worlds, or fictions? One might think, for instance, that a kind of 'ersatzist' conception of the models could help here. This is not the place to dive into an in-depth discussion of ersatzism[131], but it is clear that such a solution will depend on an ersatzist notion of a possible world (i.e., a representation of how the world could possibly be), which is known to be subject to consistency constraints: only maximally consistent representations represent ways the world could possibly be. But consistency is a logical property, just like logical truth and logical consequence, and therefore the ersatzist alternative does not help the reductive Tarskian.[132] Abandoning the conception of models based on actuality puts considerable pressure on the reduction of logical truth and consequence to material truth that is crucial for the Tarskian approach.[133]

We conclude, with Etchemendy, that the Tarskian approach is wrongheaded from the start: it makes logical consequence depend on how things *actually* are, and hence is bound to run into trouble where it comes to metaphysical matters like the size of the universe. Even if all doubt as to the extensional adequacy of the model-theoretic notion can be discarded (e.g., by embracing the necessary existence of mathematical objects and arguing that any model one might require is isomorphic to some mathematical structure or other), it is still strange to define logical consequence by reference to extra-logical matters of fact—the order of explanation is wrong, as we will see in a moment.

To be sure, model theory *as such* is strictly independent from the Tarskian project of reducing logical consequence to semantic validity. If we drop the Tarskian approach, and hence stop looking for a suitable reduction of logical consequence and logical truth to material truth, we are still free to use model theory. We then understand the foundations of model theory to be subject to modal constraints: only those models

[131] But see Fine [2002] and Merricks [2003] for critical discussion.

[132] Notice that these kinds of considerations led David Lewis to embrace his notorious thesis of 'modal realism' [see Lewis 1986a]. There are structural similarities between his reductive approach to metaphysical modality and the Tarskian, reductive understanding of logical truth and consequence we are considering—both aim to reduce modality to generality. See also fn. 124 on p. 91 above.

[133] One could, perhaps, evade some of the problems by following Williamson [2013] in adopting 'necessitism', the view that, necessarily, everything necessarily exists. It is clear, however, that this would be a rather high price to pay in order to save the Tarskian approach. Interestingly, Williamson does hint towards support for his view from Tarski's account of logical consequence and truth—see, e.g., Williamson [ibid, p. 93].

should be allowed that are logically possible. And this leaves room, of course, for disagreements as to what *is* logically possible, like those we actually find between classical logicians, constructivists, dialetheists, advocates of free logic or multi-valued logic, etc. This logical modality is one of the topics we wish to concentrate on in this chapter—without, however, deciding one way or the other when it comes to the mentioned disagreements.

Etchemendy calls such a non-reductive understanding of model theory 'representational semantics' as opposed to the Tarskian, reductive understanding, which he calls 'interpretational semantics'.[134] It is clear that we follow Etchemendy in favoring this representational stance over the interpretational stance. Now, John MacFarlane understands these two stances as taking models, on the one hand, as 'possible worlds' (representational semantics), and, on the other hand, as 'possible meanings' (interpretational semantics). His reservations concerning the interpretational stance coincide largely with Etchemendy's (and hence with ours), but he also has serious reservations concerning the representational stance. He writes:

> On the "possible worlds" approach, the domain of an interpretation represents the collection of objects that exist at some possible world. But this does not seem right, for several reasons. [. . .]
> [First:] When we take the intended domain of a theory to be, say, the set of integers, we are not engaging in counterfactual reasoning and asking what would be true in a world where only the integers exist: we are reasoning about the integers in the actual world. [. . .]
> [Second:] if domains represent the objects existing in different possible worlds, then our standard practice of allowing any set to count as the domain of an interpretation [. . .] depends on the assumption that given any set of objects, the world could have contained just those objects. This is a substantial and contentious assumption on which to base one's logical theory: it implies, for instance, that mathematical objects are not necessary existents. [. . .]
> [Third:] this approach does not make good sense of the analogous use, in modal logic, of different frames or model structures (collections of possible worlds with accessibility relations) in different interpretations. [MacFarlane 2000, pp. 196–7].

We contend that all of these objections rest on an unfortunate confusion of *logical* possibility with *metaphysical* possibility. It is ill-advised to construe models as 'possible worlds' because this seems to imply that these 'worlds' are *metaphysically* possible. Instead, we should construe the models as being *logically* possible, where we understand this type of possibility to extend beyond metaphysical possibility.

Indeed, we may even use MacFarlane's reservations to support our own quest. For we started out with the observation that logical consequence is a restriction on the

[134] See Etchemendy [1990], in particular chapter 2 for representational semantics and chapter 4 for interpretational semantics.

intuitive and more general consequence relation, on the very idea of something's following from something, and we voiced the hypothesis that this root notion coincides with the notion of conceptual consequence we aim to develop below. Hence there are bound to be logical possibilities that are not conceptually possible—which, translated into models or 'possible worlds', will amount to logically possible 'worlds' that are not conceptually possible. And, given that metaphysical modality just is conceptual modality, on our view, these worlds will not be metaphysically possible either. But let us not get ahead of things.

3.2.2 Explanatory Adequacy and Logical Necessity

We observed that our intuitive notion of valid inference, the idea of something following from something, contains a modal element: if the premises are true, the conclusion *must* also be true. This modality is entirely absent from the model-theoretic notion of logical consequence. It merely states that, under all reinterpretations of non-logical constants, if the premises are true, the conclusion *is* true. The modality is replaced by generality.

Yet even Tarski, in his original paper, makes the following remarkable claim:

> It can be proved, on the basis of this definition, that every consequence of true sentences *must* be true [Tarski 1936, p. 417, emphasis added].

Where does this *must* come from? Etchemendy accuses Tarski of committing a modal fallacy when making this claim, which he calls 'Tarski's Fallacy', of roughly the following shape. What Tarski appears to claim is this: for any logical consequence A (on his semantic definition) of a given set of premises $\mathcal{A}$, it is the case that A *must be* [true provided that all of $\mathcal{A}$ are true]. Assuming that $\varphi = \overline{A}$ and $\Gamma = \overline{\mathcal{A}}$, based on a suitable choice of logical constants $\mathfrak{F}$, we may roughly formalize that claim as follows (where $\Rightarrow$ codes the proof Tarski is gesturing towards, and $\rightarrow$ expresses purely factual truth-preservation, i.e., material implication):

$$(\Gamma \models \varphi) \;\Rightarrow\; \Box(\mathcal{A} \rightarrow A)$$

The fallacy, says Etchemendy, lies in the scope of the $\Box$-operator. It is not that, on the model-theoretic definition, A *must* be true if all of $\mathcal{A}$ are true—that modal element was eliminated, as we have seen. Rather, the most Tarski can say is that it *must* be so that *if* φ is a logical consequence of Γ, *then* A is true if all of $\mathcal{A}$ are true. That is, in our

quasi-formal speak:

$$\Box[(\Gamma \models \varphi) \Rightarrow (\mathcal{A} \rightarrow A)]$$

The fallacy Etchemendy accuses Tarski of is thus a scope fallacy.[135] Observe, however, that Etchemendy apparently agrees with Tarski that some modal force enters the scene here, although its scope differs from what Tarski seems to say. So, to repeat, we may ask: where does it come from? The answer can be found in the very same quote from Tarski: 'It can be *proved* that . . . '. The modal force in question comes with the *guarantee* that Tarski's promised proof yields. This proof *guarantees* that if the model-theoretic consequence relation holds, then the conclusion is true if the premises are true.[136] The proof brings in the modal force, which is that of the 'real' consequence relation, *not* that of the model-theoretically defined relation of logical consequence. Using our symbol $\triangleright$, which stands for the intuitive consequence relation, we may thus rewrite Etchemendy's correction of Tarski's claim as follows:

$$(\Gamma \models \varphi) \triangleright (\mathcal{A} \rightarrow A)$$

Etchemendy's bold accusation towards Tarski gave rise to quite some reactions in defense of Tarski. For instance, Ray [1996] claims that Tarski never wanted to draw any modal conclusions, that he never intended to capture any putative modal aspects of the intuitive consequence relation—and indeed, Tarski's desiderata were merely truth-preservation and formality, as we have seen. Ray suggests that Etchemendy's reading of Tarski amounts to holding that logical consequence is, in fact, strict implication, which, Ray thinks, is not a formal relation at all (we return to this charge below, in §3.3). But even Ray concedes that the 'must' in our Tarski-quote above indicates a certain modal force, one of 'logico-deductive nature'—'logical necessity, if you will' [ibid, pp. 653–4]. Well, that is, more or less, what we concluded above, with Etchemendy: the only modal force we can attach to the model-theoretic consequence relation is the necessity pertaining to the inference from $\Gamma \models \varphi$ to $\mathcal{A} \rightarrow A$, and the validity of *that* inference falls outside of the scope of the model-theoretically defined

[135] For Etchemendy's exposition of 'Tarski's fallacy', see Etchemendy [1990, pp. 85–94]. Notice that, in our quasi-formal presentation of Tarski's fallacy, we ignore a mention/use complication: Γ and φ are mentioned, not used, in order to say that they stand in the relation denoted by $\models$—much like instances of Tarski's famous T-schema contain mentions of sentences of which truth is predicated. See Tarski [1935].

[136] The proof is not a very big deal. If $\Gamma \models \varphi$ holds, then, by Tarski's definition, on any reinterpretation of the non-logical constants in all of Γ and in φ, if all of Γ are true, then so is φ. Hence this will be the case for the original interpretation of the non-logical constants in $\mathcal{A}$ and A. Hence the material implication which consists of the conjunction of all of $\mathcal{A}$ as antecedent, and A as consequent, is true.

consequence relation—it is a (valid) inference *about* the definition in question.

Although the proof towards which Tarski gestures indeed does involve a species of necessity, it is of no help to the defender of the model-theoretic notion. For suppose that we define *bogus consequence* to be the relation that holds between $\mathcal{A}$ and A just if the material conditional $\mathcal{A} \rightarrow A$ is true (assuming $\mathcal{A}$ to be finite). Of course we can now provide a (trivial) proof showing that it follows with necessity from this definition that if this bogus consequence relation holds, then A will be true if all of $\mathcal{A}$ are true. Which shows that such a proof of adequacy indeed does nothing more than simply guaranteeing that the defined notion of model-theoretic (or bogus) consequence is *in fact, materially* truth-preserving.[137]

We do not want to take any stance on the historical question as to what Tarski actually wanted or thought, on whether or not he committed 'Tarski's Fallacy'. What matters for our purposes is that the *explanation* of the putative success of the model-theoretic consequence relation crucially rests on the *guarantee* that is provided by pieces of reasoning, such as Tarski's promised proof. Such guarantee is best characterized in terms of the intuitive consequence relation itself. For, in case the reasoning in question is conducted in a formal metalanguage, it rests on the logical consequence relation for that language, which may itself be subjected to Tarski's approach. But then, the adequacy of that meta-level relation of logical consequence can again only be explained by further reasoning about the meta-language. The point should be clear: on every level, formalizing the relevant bit of reasoning invites further bits of reasoning, since for every formally characterized relation of logical consequence one may ask whether it is adequate—that is, whether it accords with the *intuitive* consequence relation. Hence it will not do to claim that the relation of logical consequence can be reductively explained in terms of generality in the way the Tarskian approach proposes. In particular, then, such a reductive approach is of no help if one aims to better understand the very idea of something's following from something, i.e., the intuitive consequence relation itself.[138]

3.2.3 Some Remarks on Logical Constants

Our assessment of the Tarskian, reductive approach to logical consequence has resulted in a firm rejection. And we have arrived at this rejection without even touching upon the demarcation question: which are the logical constants? It is instructive,

[137] See Etchemendy [1990, p. 90].
[138] Kreisel argued similarly in his [1967, see esp. §2]. See also p. 104 below.

however, to briefly consider this issue before moving on to our own construal of logical consequence in terms of a prior notion of conceptual consequence to be developed below. In particular, we are here interested in proposed solutions to the demarcation question insofar as they are taken to contribute to a reductive view on logic.[139]

We should remark that it is of course very well possible to *stipulate* that 'logic' is the study of first-order quantification theory, or something such. That is not the kind of demarcation we are interested in here: the thought is, rather, that the logical constants are marked off from other constants by having certain special features—in particular, features that can be singled out independently from logical notions. The leading thought is, presumably, that logic should be *topic-neutral*, that it provides a framework for talking and thinking about anything whatsoever.[140] That is to say, the logical constants are those with which we are not (yet) saying anything about anything in particular.

One straightforward way of spelling out this thought is in a grammatical way: the terms that belong to grammatical categories suited to construct complex sentences out of given atomic ones are topic-neutral in the required sense, so the thought goes, because it is precisely the atomic sentences that make the complex ones built out of them be about whatever they are about.[141] This proposal only makes sense as long as the relevant terms of the language are well-chosen—that is, if we make sure that only *logical constants* are used to construct complex formulae out of atomic ones. Consider deviant connectives and quantifiers such as the following:

(#) #A is true just if A is not true and Rembrandt is a painter

(&) &xA is true just if for all x, A holds and Rembrandt is a painter

These cannot be ruled out on *grammatical* grounds alone. And it is clear that similarly deviant terms can be cooked up for every grammatical category.[142]

A more sophisticated version of basically the same idea has been defended by Tarski, who indeed left open the demarcation question in his 1936 article, but revisited the issue in later work. He suggests a criterion for logical constant-hood in terms of permutation invariance—much in line with the overall picture that the Tarskian

[139] We make grateful use, in what follows, of the overview over the debate on logical constants provided by MacFarlane [2009].

[140] This thought has been voiced by, e.g., Ryle [1954, p. 116], Peacocke [1976, p. 229], and McCarthy [1981, p. 504].

[141] See, e.g., Quine [1970] and Dummett [1973]. The thought can be traced back to the distinction between 'categorematic' and 'syncategorematic' terms used in pre-Fregean (syllogistic) logic.

[142] Davidson's [1973, p. 71] proposal to take the iterative clauses (as opposed to the base clauses) of his famous truth-theoretic semantics to isolate the logical constants founders on similar issues—see, e.g., Evans [1976].

approach paints.[143] The thought, here, is that one may assign extensions on a given domain not only to individual constants and predicates of various arities, but also to sentential connectives and to quantifiers (see, e.g., McGee [1996, p. 569]). Then, one may consider which of *all* of a language's terms retain the same extension on all permutations of the domain (where permutations are bijections of the domain onto itself).

Depending on the details, the permutation invariance criterion yields more or less (quite literally) the traditional logical constants: truth-functional connectives, pure quantifiers (of different orders), and the identity predicate. It also includes predicates like 'is a thing', which invariably apply to everything, and 'is nothing', which invariably apply to nothing. However, as MacFarlane [2009, §5] observes, this view has some strange consequences if we stick to an orthodox, interpretational understanding of what is included in the domain that we permute, that is, if we assume the domain to contain only what actually exists. For instance, if it so happens that no two things have exactly the same mass, then the relation of having exactly the same mass will have the same extension as the identity relation and hence will be counted as a logical constant likewise. Moreover, the deviant connective # and quantifier & given above will come out as logical constants too.[144]

The problem is that contingent matters of fact are interfering with the demarcation of logical constants on the proposal at hand—much like such contingent matters of fact influenced the extension of the orthodox, Tarskian definition of logical consequence (see §3.2.1 above). A way out, then, could be to consider permutation invariance not only in the actual domain, but in all possible domains. But, as we remarked earlier (see p. 95 above), the notion of possibility involved at least requires consistency—which makes trouble for the proposal.

In effect, invoking possible domains only shifts the wrinkle in the carpet. For if the sense of 'possibility' used is metaphysical, there will be metaphysical truths that are not supposed to be logical truths (e.g., water is H_2O, or other such 'a posteriori necessities') that can be encoded into a quantifier or connective (yielding, for instance, 'there exists an x such that water is H_2O and $\ldots x \ldots$' or '$\neg \ldots$ and water is H_2O'). And if the sense of 'possibility' used is epistemic, then there will be analytic truths that are not supposed to be logical truths (e.g., bachelors are unmarried) that can be used to similar effect. Such deviant 'logical constants' can only be ruled out if the

[143] See Tarski [1986], and also McCarthy [1981], van Benthem [1989], Sher [1991, 1996, 2003], McGee [1996], and Feferman [1999].

[144] See McCarthy [1981, §V] and McGee [1996, p. 575] for similar arguments.

sense of 'possibility' in question is *logical possibility*—but it is clear that we are then moving in a circle.[145]

Leaving such a semantical approach aside entirely for a moment, we may consider a different proposal: identifying the peculiar inferential properties that logical constants have (and other terms lack). On this 'inferentialist' or proof-theoretic alternative, the detour through models or other semantic machinery is unnecessary. For if the logical constants can be characterized by providing the relevant rules of inference (introduction and elimination rules, say), and these rules can be shown to be special in *merely* providing ways to infer new statements from given statements, independently of what the statements are about, it seems that we have found a precise way of identifying the logical constants as the topic-neutral constants in just that sense.[146]

Such inferential or proof-theoretic approaches to understanding logical constants provide interesting explications of their significance (just as semantic approaches do), but they will not serve the reductive demarcator's purpose, as Prior's [1960] famous example of a propositional connective 'tonk' makes clear: the introduction and elimination rules for 'tonk' are indeed purely inferential, but in effect they allow inferring anything from anything.[147] The constraints required to exclude such deviant inference rules (not to mention rules for the deviant, extra-content-invoking quantifiers and connectives we considered earlier) in effect come down to assumptions that threaten the reductive proof-theoretic demarcation project (but not the project of explication) with circularity.[148] Indeed, it is not uncommon for defenders of this inferential characterization view on logical constants to hold that the basic inference rules should be 'primitively compelling' (see, e.g., Peacocke [1987, p. 156], Sainsbury [2001, pp. 370–1], and Hodes [2004, p. 136–9]), which is to say that they should amount to basic steps of reasoning the intuitive validity of which is beyond doubt.[149] It should not be the aim of such inferential characterizations of logical con-

[145] In fact, there will even be deviant connectives if the sense of possibility used is logical possibility, for one can then encode logical truths into the sense of putative connectives and quantifiers.

[146] See, e.g., Gentzen [1969], Hacking [1979], Prawitz [1985], Peacocke [1987], Dummett [1991], and Hodes [2004].

[147] 'Tonk' is governed by an introduction rule akin to that for the disjunction, and by an elimination rule akin to that for the conjunction.

[148] For instance, Hacking [1979, pp. 311–14] argues that, assuming bivalence, proof-theoretic constraints can be given for introduction and elimination rules in a sequent calculus that ensure the logical constants thus defined to behave properly (these constraints include the subformula property and the provability of elimination theorems for Cut, Identity, and Weakening). And Hodes [2004, §IX] argues that only pairs of elimination and introduction rules that are 'maximal inverters' of each other and together determine a unique sense constitute the sense of a true logical constant.

[149] This aspect of the inferentialist approach clearly has an intuitionist (or perhaps Kantian) ring: it is

stants to provide non-circular demarcations of the logical constants or to contribute to reductive accounts to logical consequence.

To conclude, both the proof-theoretic (inferential) and the model-theoretic (semantic) approach to understanding logical constants, once divorced from any reductive ambitions, may be viewed as valuable sources of insight into the significance of those logical constants—for instance, they allow one to locate in a precise way the differences between, e.g., intuitionist, classical, and paraconsistent readings of certain logical constants. For our purposes, it is important to keep in mind that, properly understood, such results do not demarcate the realm of logic proper but are rather to be read as explicating the conceptual relations that hold within certain families of concepts, such as certain families of propositional connectives (e.g., the truth-functional ones), or the family of concepts belonging to first- or higher-order quantification theory, or certain families of modal or temporal or deontic concepts, etc. The accompanying relations of logical consequence (both model-theoretically defined and proof-theoretically defined) may capture the intuitive consequence relation $\triangleright$ *insofar as the concepts in question are concerned*. This will form the core part of our own understanding of logic to be developed below.

3.2.4 Concluding Remarks on the Tarskian Approach

The model-theoretic construal of logical consequence is only of interest if we consider it non-reductively as coinciding with the class of intuitively valid arguments within a certain well-defined area of logical study (first-order predicate logic, say). Luckily, we have a completeness theorem for first-order logic, which guarantees that our model theory for first-order logic adequately captures only first-order logical possibilities—completeness guarantees that for every model-theoretic consequence φ of a given set Γ there is also a proof of φ from Γ. It is the latter, constructed as it is with the help of *intuitively* sound axioms and/or rules of inference, which can carry the required modal force. That modal force derives from the intuitive relation of consequence $\triangleright$.

We have excerpted from Etchemendy's work two main reasons for rejecting the reductive, Tarskian approach to logical consequence. First, we have seen reasons to doubt the extensional adequacy of the model-theoretic consequence relation $\models$ (in §3.2.1 above), given that it is affected by how things actually are. And second, we

built, in effect, on basic logical intuitions. (Even if these are, subsequently, 'naturalized'.) Notice that we should not think of 'primitively compelling' in merely subjective terms: as Martin-Löf [1987] argues, there is a normative notion of correctness at play here which distinguishes the intuitionist standpoint from a mere subjectivism or relativism.

noticed that, in order to show the model-theoretic conception of logical consequence to be any good, you have to infer something *about* it (§3.2.2). There is no escaping the intuitive consequence relation $\triangleright$, with its typical modal force. For instance, there is no *guaranteed* truth-preservingness without a *guarantee of* truth-preservingness—a proof.

The very project of showing that a candidate reductive definition of logical consequence is adequate presupposes the intuitive notion of consequence. Likewise, the project of isolating all and only those terms that deserve the status of being logical constants (which we briefly surveyed in §3.2.3 above) presupposes an intuitive grasp of what the logical constants are, or, equivalently, of a specifically logical notion of possibility. Thus we are brought back to the proper focus of the present chapter: the intuitive consequence relation $\triangleright$. No elucidation of that idea is forthcoming from attempts at reduction like the Tarskian approach—nor, for that matter, from alternative, non-semantic reductive approaches. That is, we agree with Kreisel, who writes:

> Nobody will deny that one knows more about [$\triangleright$] after one has established its relations with [$\models$] and [$\vdash$]; but that doesn't mean that [$\triangleright$] was vague before. [Kreisel 1967, p. 154; notationally adapted]

Kreisel means to resist the idea that we should somehow replace the intuitive notion with the formally defined notions: we assess the adequacy of the latter by means of the former.

In our considerations so far, we have largely taken for granted one quite important dimension of the specifically logical consequence relation: its being a *formal* notion. Now that we have abandoned the search for a reductive demarcation of logical constants as well as the attempt to arrive at a reductive understanding of logical consequence, we may still ask what it is about logic that makes it formal, and how that notion of formality relates to the intuitive consequence relation, to the very idea of something's following from something. Our most important task will be to connect the required notion of formality with the notion of (logical) necessity and possibility that has resurfaced throughout our above discussions.

3.3 Truth Guaranteed By Content

Our assessment of the Tarskian, reductive approach to logical consequence has resulted in a firm rejection of such approaches. But it also raised a number of interesting questions concerning the nature of and relation between logical form and logical necessity.

Interestingly, the question of how formality and necessity relate has cropped up in the small literature spawned by Etchemendy's work. Several of his critics complain that, in accusing Tarski of committing 'Tarski's Fallacy', Etchemendy confuses valid inference with strict implication.[150] But, these critics contend, *formal* consequences, which is what logic is concerned with, are not to be identified with *necessary* consequences (even if, as is natural to assume, all formal consequences are necessary consequences too). Two examples:

> The relation Tarski wants to capture is a *formal* relation. ... However, strict implication is not a formal relation in this sense [i.e., uniquely determined by the form of the sentences between which it holds]. Hence, Tarski's target notion *is not strict implication*. [Ray 1996, p. 650–1, emphasis in original]

> If A is a logical consequence of $[\mathcal{A}]$, then it is presumed that A must be a *necessary* consequence of $\mathcal{A}$. But most logicians do not recognize all the necessary consequences of a sentence as logical consequences. [Warmbröd 1999, p. 513, notationally adapted]

These statements are presented as reminders of what we—supposedly—all very well know: logical consequence is a purely formal matter, whereas necessity is not. Whether or not an argument is logically valid depends on the logical form of its premises and conclusion. As we saw, logical form is traditionally thought to be determined by the set $\mathfrak{F}$ of logical constants: the logical form of a given sentence A, which we denote by $\overline{A}$, consists of the occurrences of logical constants together with the pattern of occurrences of non-logical constants (roughly).

The problem with these critical responses to Etchemendy's conclusions is that they, again, trade on an important ambiguity in the terms 'necessity' and 'strict implication'. There are various notions of necessity and possibility, and to each of these there corresponds a separate notion of strict implication. In particular, the notion of necessity that Etchemendy is interested in is *logical* necessity, which supposedly characterizes the relation of logical consequence. And it is not at all clear whether the 'formal consequences' are different from the *logically* 'necessary consequences' at all—indeed, once the ambiguity is dissolved, it seems rather natural to identify logical consequence with *logical* strict implication. If the premises of a logically valid argument are true, it is logically impossible that the conclusion is false. What is potentially confusing here is the strong association we tend to have between modal notions in general and the various systems of modal logic—we are not assuming in advance that any of the relations of strict implication familiar from

[150] Strict implication is a relation that holds between φ and ψ just if $\Box(\varphi \to \psi)$. Which understanding of the $\Box$ is involved is usually left open—and this makes for important differences; see below.

such systems coincides with logical validity.[151]

So, formality and necessity are not as independent as one might think, at least as long as we take 'necessity' to mean logical necessity. That is in line with several observations we have made so far—for instance, that a proper model theory has to make sure that the range of suitable models coincides with the range of logical possibilities.

3.3.1 Reconstruing Logical Form

It is time to move on to our own construal of logical form. That is, we will first, in a way, complete our deconstruction of the notion, in order to identify its conceptual core, which will then guide us in construing it differently, leading to a notion of logical form that is useful for our project of explicating the Aristotelian picture.

We take our cue, again, from Tarski:

> [S]ince we are concerned here with the concept of logical, i.e. *formal*, consequence, ... this relation cannot be influenced in any way by empirical knowledge, and in particular by knowledge of the objects to which the sentence [A] or the sentences of the class [$\mathcal{A}$] refer. [Tarski 1936, pp. 414–5, emphasis in original]

Tarski's remark on formality is best understood against the background of the thought that logic is topic-neutral (see §3.2.3 above). It seems that Tarski's take on this thought is a fairly traditional empiricist one: formality is connected to analyticity and to aprioricity, to being independent of empirical information. Consequently, form is *opposed* to content: form merely provides the scaffolding into which content—empirical information—can be poured in various ways to yield truths or falsehoods—and in some cases, the content doesn't make any difference as to the truth or falsehood of the resulting sentence: those are the logical truths.[152]

We highlight two crucial thoughts in this background empiricist view which also play an important motivating role in the Tarskian approach. The first is the thought that form is what you get when you remove all content, such that form is the absolute *absence* of content. The second is the thought that forms do, in some cases, *all by themselves* guarantee the truth of their instances. We challenge the first thought, but embrace the second, in a qualified way.

[151] Though we should note that Carnap, for instance, developed systems of modal logic that were quite explicitly meant to capture the logical modalities—i.e., such that $\Box\varphi$, in such a system, holds just if φ is a logical truth. Arguably, then, strict implication just is logical consequence in such systems. See Carnap [1946, 1956/1947]. Williamson [2013, ch. 2] provides a nice summary of Carnap's logical approach to logical modality.

[152] A paradigmatic exposition of this kind of view, insofar as logical modality is concerned, was developed by Carnap [1946, 1956/1947, 1950].

The idea that form is what is left over if you leave out *all* content is, on the face of it, quite puzzling. For how could such a form guarantee truth? There is nothing in the *typographical* form of, say, $Pr \rightarrow Pr$ which could possibly guarantee its truth. What is crucial is the *contribution* that the logical constant $\rightarrow$ makes to the *content* of the entire sentence $Pr \rightarrow Pr$ (amongst other things). Hence form cannot be that which is left over if you remove *all* content: that would leave us with nothing but mere scribbles on paper or sounds in the air. Apparently, then, it is what is left over if you *partially* remove the content—in particular, if you remove all but the content that is held fixed in $\mathfrak{F}$ (and the pattern of occurrences of terms not included in $\mathfrak{F}$).

As we saw, Tarski's strategy of eliminating those aspects of the content not in $\mathfrak{F}$ was, first, to consider arbitrary substitutions of all non-logical expressions, but because of technical problems, that strategy was then refined in terms of arbitrary reinterpretations (i.e., in terms of satisfaction). On closer inspection, however, this is quite a strange approach: if logical truth and logical validity are indeed purely formal notions, then it makes much more sense to focus on what logical forms *do* instead of focusing on what they *don't* do, so to speak—after all, it is supposed to be the logical forms in virtue of which certain arguments are logically valid and in virtue of which logical truths are true. To put it less mildly: the question what happens under arbitrary substitutions of all non-logical constants is as irrelevant to logical consequence as is the question what happens under arbitrary substitutions of, e.g., typographical fonts in which the relevant formulae are written down. In both cases, we learn *that* logical consequence and logical truth is stable under the relevant substitutions—not *why* that is the case.

Let us see whether we can make sense of truth and validity in virtue of form without recourse to either substitution or reinterpretation. As our point of departure, we borrow an interesting thought experiment from Kit Fine: imagine a device that takes a contentful sentence A as input.[153] Let us say that the device is equipped with a filter through which it can only 'see' the logical form of its input sentence, $\overline{A}$, nothing else: if the device can decide that its input sentence A is true simply on that basis, then it is logically true.[154]

Notice that our imagined device need not look at any models: the logical form suffices. That is to say, the device is able, we are supposing, to see what contribution

[153] We ignore logical consequence and focus on logical truth for a while. The conceptual points we will be making will carry over unproblematically.

[154] Fine uses the idea of such a device to illustrate his notion of 'logical satisfaction'. In Fine's version, the device takes a (possibly open) formula *and* a sequence of objects as input—for our purposes, this complication is not necessary; Fine considers issues related to but not the same as ours. See Fine [1989, pp. 46–7].

the logical form $\overline{A}$ makes to the overall content of the input sentence A, and is able to decide on the basis of that information whether or not the sentence is logically true. Notice that our device minimally requires that the input sentence is indeed a contentful sentence, hence aspects of it that are filtered out (i.e., that occur in A but not in $\overline{A}$) are still required to be meaningful.[155] In Fine's [2005b, §§1–2] terms, then, the sentences that our device judges logically true are *transcendental*: they are true 'regardless of the circumstances', not merely true in every circumstance (which is what Tarski's model-theoretic definition in effect rests on). Fine's thought, here, is that necessary truths are true in all possible worlds whereas transcendental truths are true independently of the possible worlds—much like eternal truths are true at all times while atemporal truths are true independently of time (Fine calls these latter truths 'sempiternal'). This notion of transcendentality, of independence of how things are, is of interest for us. However, we will shortly find reason to depart from Fine's way of distinguishing it from necessity.

Let us consider the question what our device could possibly take as input if it is to function in the described way, that is, without having to look beyond the content it is presented with. If the device can judge its input sentence to be true (or false) on the basis of the parts of its content it is allowed to 'see', the resulting judgment will be a transcendental matter in the Finean sense. We should keep in mind our conclusions with regard to the various approaches to understanding logical constants above: we decided that they are best viewed as explicating the significance of logical constants, e.g., in a model-theoretic or proof-theoretic way. Now, we can envisage many different types of filters, so to speak, that allow our device to discern more or less aspects of the significance of such logical constants, and even aspects of the further content of the sentences it is presented with. Different filters will then result in different outcomes for the same sentences. For suppose we present our device with the following input sentences:

(1) If Rembrandt is a painter, then Rembrandt is a painter.

(2) If Rembrandt is a painter, then someone is a painter.

(3) If Rembrandt is a painter, then Rembrandt is an artist.

(4) If Rembrandt is a painter, then Rembrandt is mortal.

(x) If Rembrandt is a painter, then Rembrandt has brown hair.

[155] Thus a sentence like 'All mimsy borogoves are mimsy borogoves' only comes out as logically true (given a predicate-logical filter, see below) if the term 'mimsy borogove' has content—otherwise, we may assume, the device returns an error.

(1) will be judged true by our device if the filter we insert into it is a filter of propositional logic, i.e., if it would only let the device 'see' the truth-functional aspect of the contributions that sentential connectives make to the content of the presented sentence, as well as the pattern of the atomic propositions they operate on. For (2) to be judged true by our device, we would have to replace that filter with a predicate-logical one which allows the device to also recognize quantifiers and the pattern of occurrences of predicates of various arities and of individual constants. For (3) to be judged true by our device, we should insert a filter which allows the device to see aspects of the contribution the predicates themselves make to the contents of the sentences. Analogously, (4) will be judged true if we insert a filter which allows the device to look at the (ontologically fundamental) kinds of objects named by individual constants (viz., that 'Rembrandt' names a human being). But for (x), of course, no filter can be given that will allow our device to judge it true: its truth cannot be settled even if we keep all content in place. Rembrandt's hair color is a contingent matter to be decided by empirical investigation, after all.

Observe, now, that *all* of the presented cases, except for (x), are completely analogous insofar as the guaranteeing of truth is concerned: truth is guaranteed by parts of the overall content of the sentences presented. They are all true, then, regardless of the circumstances—and that makes them transcendental truths, to borrow Fine's label. We should stress, however, that we thereby immediately part company with Fine on the notion of transcendentality. For he writes:

> The sentence 'Socrates exists or does not exist' ... is a necessary truth; its truth will indeed depend upon the circumstance (and, in particular, upon whether or not Socrates ends up existing) but in such a way that it is true whatever the circumstances, or however things turn out.
> Of course, there is a sense in which the sentence 'Socrates exists or does not exist' is true regardless of the circumstances, for we can recognize it to be true on the basis of its logical form alone and without regard to the circumstances. But this is not the sense of 'regardless of the circumstances' that I have in mind. We should imagine ourselves evaluating the truth-value of the sentence on the basis of its logical form. This means that we should first evaluate the disjuncts 'Socrates exists' and 'Socrates does not exist' in turn and see what their truth or falsehood might depend upon; and since they depend for the truth upon the circumstances, then so does the disjunction. [Fine 2005b, p. 324]

In effect, then, we are operating with the alternative notion of transcendentality that Fine recognizes but puts aside—which is fine, of course. The image of our device (which, to repeat, we also borrowed from Fine) compels us to adopt the notion Fine here dismisses, and to dismiss the notion he adopts: we take our device to be blind to the aspects of the presented content that are filtered out, so to speak, and therefore

to be blind to whether or not these ignored aspects of content depend, for their significance, upon the circumstances. In our examples (1)–(4), the 'Rembrandt is a painter'-bit that occurs in all of them will depend, for its truth, on the circumstances, of course. Hence (1)–(4) are not transcendental truths in Fine's sense. But our device is blind to that information, yet still able to decide the sentences to be true (given the right filters). Hence they are transcendental truths in our sense (which is the alternative sense Fine dismisses). Compare (x): its truth crucially depends upon the circumstances, and in particular, upon whether or not Rembrandt has brown hair. In such cases, truth is not guaranteed by content alone but depends upon how things are. Thus we arrive at precisely the notion of truth guaranteed by (parts of) content we wish to isolate.

The image of our device can help us understand what makes the demarcation project for logical constants so puzzling—as we observed in §3.2.3 above. For a start, instead of asking which terms should be included in $\mathfrak{F}$, the question now becomes: which of the many different filters is the genuinely 'logical' one? This way of putting the question is helpful because it focuses on the *kinds* of contributions to content instead of focusing on the *full* contribution that specific terms, constants, or symbols make. Recall the sentential operator #, which we defined as follows:

(#) #A is true just if A is not true and Rembrandt is a painter

makes a complex contribution to statements in which it occurs: it operates on its argument in the way ¬ does and also contributes, in conjunction with the result, the content presented by 'Rembrandt is a painter'. When equipped with a propositional-logical filter, we are supposing our device to be able to discern the relevant aspects of this contribution. Hence it will judge the following sentences logically true:

(5) If #(snow is black), then snow is not black.

(6) If #(snow is black), then Rembrandt is a painter.

Equipped with a filter which also takes the contribution made by predicates into account, our device will also recognize the following sentence to be true:

(7) If #(snow is black), then Rembrandt is an artist.

It is much more fruitful, then, to think in terms of *kinds* of contributions to content than to think in terms of a set $\mathfrak{F}$ of logical constants which are to be held fixed. This thought will play an important role below.

Frege taught us to look beyond the surface structure of our natural languages, to which the traditional syllogistic understanding of logic had been faithful for centuries, and showed us that logical form does not always coincide with grammatical form. This insight allowed him to develop his great contribution to the discipline of logic: isolating certain specific aspects of content and developing a logical system for representing and investigating these aspects. The result was predicate logic as we now know it, in which certain kinds of contributions to content are isolated in a syntactically tractable way, i.e., by way of the familiar logical constants (including the form of predication involved, we should add).[156] That is to say that connectives like # do not occur in this system. It is far from obvious, however, that *every* aspect of content can be spelled out in terms of just those aspects that Frege uncovered—we discuss this question below (§3.3.2).

Looking again at our list of input sentences (1)–(4), one can say that, traditionally, (1) and (2) are considered *logical* truths, while (3) and (4) are considered true in virtue of meaning, that is, *analytically* true. We can, of course, supplement truths like (3) with further information in order to arrive at truths which are logically true in the traditional sense (e.g., by adding 'and all painters are artists' to the antecedent of (3)). But it is not self-evident that that is an adequate way of explicating the relevant aspect of content—after all, the truth of 'all painters are artists' is *itself* guaranteed by its content. And, more generally, there is no guarantee whatsoever that such a predicate-logical rendering will be available for *all* sentences that our device may judge true on the basis of more or less strict filters, even if we set aside modal notions. Be that as it may, our considerations so far compel us to conclude that all of these truths are true 'in virtue of meaning', or, as we say, their content guarantees their truth. *As far as this notion of truth guaranteed by content is concerned*, there seems to be no interesting non-arbitrary distinction between form and content, no principled demarcation of specifically logical form as that which guarantees truth in a particularly logical way.

Now, taking the intuitive consequence relation into account again, we conclude that the core phenomenon of something following from something rests on the phenomenon we have been trying to bring to the fore: that of truth being guaranteed by content *alone*. The intuitive consequence relation holds whenever the *entire* content of the relevant premises and conclusion is such that it guarantees the truth of the conclusion conditional on the truth of the premises. Restriction to certain particular aspects of that content is just a way of isolating some area of logical (or conceptual) study. What is traditionally considered to be 'logical form', which was greatly

[156] See Frege [1879, 1892a,b].

clarified and precisified by Frege, is, thus, not something definite to which content is opposed, but rather one of the many possible restrictions on the broader notion of truth guaranteed by content. Obviously, the traditional demarcation(s) of logical form are useful, since they allow for the systematic study of some very basic aspects of content. But, again, what makes *logically* valid arguments valid isn't any different from what makes valid arguments valid in general.

We conclude that the intuitive consequence relation coincides with what we may now call the relation of *conceptual consequence*: intuitively valid arguments are conceptually valid arguments.[157] And, similarly, the notion of truth guaranteed by content, on which this notion of conceptual consequence or validity rests, coincides with the notion of *conceptual truth*. Finally, recalling the close link between logical modality and logical truth and consequence, we identify the modality that characterizes the notions of conceptual consequence and conceptual truth as *conceptual modality*. The natural thought underlying these identifications is that content is *ipso facto* conceptual, that concepts are the building blocks of content.[158]

In MacFarlane's [2009] terminology, we have now positioned ourselves near the camp of the 'debunkers', who hold that the problem of demarcating the set of logical constants is a pseudo-problem. MacFarlane writes:

> [Debunkers] do not dispute that logicians have traditionally concerned themselves with argument forms in which a limited number of expressions occur essentially. What they deny is that these expressions and argument forms define the *subject matter* of logic. On their view, logic is concerned with validity *simpliciter*, not just validity that holds in virtue of a limited set of "logical forms". ... What makes people logicians is not their concern with "and", "or" and "not", but their concern with validity, consequence, consistency, and proof ... [MacFarlane 2009, §8].

This nicely summarizes the negative upshot of our view. Of course, insofar as there are reasonable grounds for restricting the kinds of content on which logicians should focus, there is nothing against fashioning corresponding notions of logical form, logical consequence, logical truth etc. There is nothing against demarcating

[157] To be sure: we do not understand the 'intuitively valid arguments' to include all and only arguments that people *in fact* find compelling, or something along these lines. We mean the 'pretheoretic' notion of validity, which is normative, not descriptive.

[158] One might worry that we are ignoring an important point of controversy, namely over whether or not there is such a thing as 'nonconceptual content' (see, e.g., the essays in Gunther [2003]; the notion originates from Evans [1982, see esp. §§5.2, 6.3, 7.4]). Strictly speaking, however, that controversy concerns certain questions in the philosophy of mind, namely whether or not there can be forms of cognition without graspings of concepts (that is, in our terms, without employing conceptions). We do not wish to take a stand on this issue in the context of this essay, though it is perhaps fair to point out that the very notion of non-conceptual content is likely to be unintelligible from the perspective of the Aristotelian picture (and we may point to McDowell [1990, 1994a,c,b, 2011] in support of this claim).

the academic discipline named 'logic' in such a way. What we wish to strongly oppose is the thought that such a demarcation can be used for a reductive account of the intuitive consequence relation—not even for a reductive account of the logical variety of that notion. Hence our sympathy for Etchemendy's critique of the Tarskian approach.

On a more positive note, we wish to suggest the following. Given the fundamental notion of conceptual truth, of truth guaranteed by content, and the possibility of restricting that notion to certain parts or aspects of content, we may cast a fresh look on the question which kinds of restrictions yield interesting fields of conceptual study. In particular, we may abandon the bottom-up project of extending traditional logical systems (first-order predicate logic, say) with additional machinery (e.g., modal operators). That bottom-up project is at least implicit in the way logic is brought to bear on questions in, e.g., metaphysics. Instead, we may look for other areas of conceptual study that might perhaps not be amenable to rigorous formalization of the familiar kinds but are nevertheless very interesting from a metaphysical point of view.[159] After all, on the Aristotelian picture, which we aim to defend in the present essay, the conceptual is not divorced from the real, and hence, conceptual investigations simply *are* metaphysical investigations.

A few orienting remarks on this last thought may help clarify the place that the investigations of this chapter occupy within the larger context of the present essay.

3.3.2 Beyond Predicate-Logical Form

Once we realize that the closely related phenomena of truth being guaranteed by content, and of something following from something, are *not* captured by any notion that can be reductively defined by methods familiar from the discipline of logic, we start to suspect that the paradigm status of certain formal systems within analytic philosophy, in particular of predicate logic, may well have done more harm than good outside of the discipline of logic. It has created a focus on particular kinds of content to the exclusion of other kinds of content—indeed, it has encouraged attempts to analyze all other kinds of content in predicate-logical terms ('regimentation' is the term familiar from Quine). Metaphysical modality is a telling case in point: typically 'logical' approaches to that notion have resulted, for the most part, in perplexing views. On the one hand, the very notion has been discarded because it

[159] Aristotle seems to have recognized such a need for diversity in the way we think about reality—he writes in the *Nicomachean Ethics*: 'it is a mark of an educated person to look in each area for only that degree of accuracy that the nature of the subject permits.' [Aristotle 2000, I.3, 1094b23–26].

supposedly resisted a sensible analysis in predicate-logical terms, as Quine's famous and long-standing skepticism about *de re* modality illustrates[160], while on the other hand analyses have been proposed that entail notoriously unbelievable but quite stubborn metaphysical outlooks, as is illustrated by Lewis's equally famous 'modal realism', but also, more recently, by Williamson's 'necessitism', the view that necessarily, everything necessarily exists.[161] Let us consider, therefore, what exactly the shortcomings of predicate logic are.

The kind of content on which predicate logic focuses is, one might say, that of predication and quantification considered *in abstracto* (in addition to the standard propositional connectives). *What* is predicated and *of what* is predicated is treated only insofar as that is required for predication to be intelligible. Frege's understanding of predication is quite salient: predicates are functions, the objects to which they apply are arguments, and from a 'logical' point of view that's all we need.[162]

Now, legitimate though the study of predication on this abstract level is, it is not very hard to see that there are *other* sorts of predication that do not fit the Fregean model (or, at least, not straightforwardly). One example, that stays fairly close to the original Fregean idea, is Laycock [2006], who argues that predication of stuffs and pluralities cannot be reduced to 'traditional' predication of individuals. A more radical example is the so-called A-theory of time, defenders of which typically insist on an irreducibly tensed form of predication. And, thirdly, we should mention M. Thompson [2008], who introduces yet another kind of predication underlying what he calls 'Aristotelian categoricals' like 'The bobcat breeds in spring', which again cannot be reduced to either traditional Fregean predication or the mentioned alternatives.[163] On reflection, it seems that the Fregean understanding of predication is particularly geared towards mathematical contexts, in which tenses, teleology, and the differences between stuffs and things play no role.[164]

But predication itself is not the only concept on which predicate logic creates a

[160] See, e.g., Quine [1943, 1947, 1948, 1953b,c, 1976b, 1991].

[161] See, respectively, Lewis [1968, 1973b, 1986a] and Williamson [1999, 2010, 2013]. Lewis's modal realism paradigmatically illustrates the orthodox project of reducing all content to what can be captured in predicate-logical terms. Williamson's necessitism is not reductive in nature—his logic for metaphysical modality properly extends predicate logic—but still suffers from the limitations that predicate logic imposes (on which more below), which play a large role in his argument for necessitism.

[162] See Frege [1879, 1892a,b].

[163] The second part of this essay largely amounts to a defense of such a pluralism about forms of predication. We provide an elaborate introduction of the forms of predication to be defended in chapter 4. In particular, the tensed form of predication is introduced in §4.3 and defended in chapter 6, and the (teleological) form of predication Thompson highlights is introduced in §4.4 and defended in chapter 7.

[164] We will devote quite some space, in the second part of this essay, on attempts to reductively define these allegedly non-Fregean forms of predication in predicate-logical terms.

bias in taking a particular stance. The same happens with the notion of a predicate. Predicate logic gave rise to the curious, purely extensional conception of properties as sets of things. All predicates are thus taken to be strictly independent from each other; to each an interpretation is given (a subset of the domain) completely independently of all others. But there is a far more natural alternative understanding of predicates according to which predicates denote properties that occupy a certain place within a range of contrarieties, a range of properties which exclude one another. For instance, having a mass of 10kg is such a property; it excludes having a mass of 9kg, having a mass of 11kg, etc. Elder [2011, ch. 8] defends just such a conception of properties.[165] Thus it will be a conceptual truth that things having one property from such a contrariety range do not have any other property from that range. Further differentiations can be made, for instance, between the features, the powers, and the causal activities of concrete things.

And, finally, the notion of an individual which is tacitly assumed by predicate logic also allows for alternative conceptions. In effect, predicate logic treats its individual constants as naming 'bare particulars' (in the sense of the previous chapter, see §2.3), as being marked off only by their self-identity. But, as we have argued in the previous chapter, sticking to this conception is only to be advised if one is willing to adopt one of the most austere ontologies there is. If one is not so inclined, other options for the treatment of individual constants arise, for instance, as denoting objects falling under one or another fundamental ontological category. Under such a conception of individuals, the referent of a particular individual constant is bound, by conceptual necessity, to have certain properties and to be related in certain ways to other things (whose existence is thus required). For instance, the existence of {Socrates} implies the existence of Socrates, and the existence of Socrates implies his being human. More generally, the fundamental category under which a thing falls may well determine what forms of predication and what kinds of predicates are applicable to it.[166]

The above sufficiently illustrates the rich variety of kinds of content available even when staying on the fairly generic level of predication as such, of 'saying something of something'. Paying close attention to this variety, we suggest, enables us to isolate distinctive kinds of content the conceptual study of which may well be worthwhile. Whether or not one is willing to call the resulting areas of conceptual study 'logic' is a terminological matter we do not wish to settle here. The second part of this essay explores three distinctive such groups of concepts, each of which comes with its own

[165] We remarked on this understanding of properties already in §1.3, p. 42.
[166] As we said, these remarks will be expanded to much greater length in the second part of this essay.

style of predication, its own conception of predicables, and its own conception of things one can predicate of.

It should be obvious, by now, that our departure from logical orthodoxy, so to speak, is quite radical.[167] It seems appropriate, therefore, to be explicit about a number of ramifications this departure has, in particular for our understanding of the traditional notions of analyticity and *a prioricity*.

3.3.3 Conceptual Truth and Empirical Truth

Consider the following quote from MacFarlane, who characterizes the position of the 'debunker', with whom we sympathize, as follows:

> [The debunker holds] that deductive validity is a feature arguments have by virtue of the meanings of the terms contained in them, so that anyone who understands the premises and conclusion of an argument must be in a position to determine, without recourse to empirical investigation, whether it is valid. On this conception, logic is the study of *analytic* truth, consequence, consistency, and validity. [MacFarlane 2009, §8]

We prefer, expectably, to replace 'analytic' by 'conceptual'. However, that is not a mere stylistic preference: we have a very specific reason for it. On reflection, there is something odd about the notion of analyticity, of truth in virtue of meaning, as it is usually understood—and here we do not mean to refer to Quine's famous attack on the notion of analyticity.[168] Rather, we suspect that the notion, as usually understood, involves an aspect of socio-epistemic contingency that makes it inappropriate as a characterization of our views. Consider the following two sentences:

(a) Bachelors are unmarried.

(b) Whales are mammals.

The first, (a), is a typical example of an analytic truth: it is true 'in virtue of meaning'. But what about (b)? Is it part of the meaning of 'whale' that whales are mammals? What makes such questions difficult to assess is that they seem to inquire into what *most people take* the meaning to be—that is the socio-epistemic aspect we wish not to have in our notion of conceptual truth. There are, presumably, still people who think

[167] For a similarly radical critical discussion of the shortcomings of first-order predicate logic, see B. Smith [2005], who argues that the exclusive focus on predicate logic's syntactic division into individuals and predicates has pervasively distorted philosophy within the analytic tradition.

[168] See Quine [1951], and for general discussion Rey [2012]. For what it is worth, we sympathize with Gutting [2009, ch. 1] that Quine's rejection of the analytic/synthetic distinction 'seems supported by nothing more than persuasive rhetoric' [p. 30].

whales are fish. Do they fail to grasp the meaning of 'whale'? Did the majority of human beings before the discovery that whales are mammals inadequately grasp the meaning of the term they used? Or did the meaning somehow change over time?

Instead of embarking on a separate investigation into the topic at hand, we simply set forth our diagnosis of the situation, which is in line with our considerations so far. The problematic aspect of the notion of 'analyticity' relates to a distinction we made, in chapter 1, between *concepts* and *conceptions*: the latter amount to more or less complete and adequate graspings of the former (see 1.3). In the second chapter, we defended a specific version of essentialism, on which the essences of things coincide with certain (ontologically fundamental) concepts. Hence, the concept of a whale, in our sense, amounts to the (general) essence of whales.[169] Insofar as people try to grasp that concept, e.g., when observing whales or reading about them in a book, they form a more or less complete and adequate *conception* of whales. Now, arguably, (a) above is analytic in the sense that the concept of whale includes their being mammals—which is to say that (a) is a conceptual truth in our sense—but not in the sense that, e.g., it follows from the conception of whales the majority of people around the year 1500 had (or, if you prefer, the conception of whales that the specialists (the biologists) living around 1500 had). Indeed, it is likely that 'mammals are fish' was analytic in this sense around 1500. Analyticity in this second sense depends on the epistemic situation of most people and is therefore unstable.[170]

Let us, therefore, call analytic truths based on *concepts* 'conceptual truths', and analytic truths based on *conceptions* 'epistemically analytic truths'. For our purposes, this second notion is not very interesting, but the first is. That is the notion we have been working towards in this chapter: truths whose truth is guaranteed by content alone are conceptual truths in precisely this sense.

Observe that our understanding of conceptual truth has some striking consequences. First, we may be *wrong* about which are the conceptual truths, simply because we have failed to adequately grasp the concepts in question. Second, and relatedly, we may be right about certain conceptual truths but fail to notice that we have grasped the wrong concepts (as in, e.g., 'witches have magic powers' while thinking of certain people one takes to be witches). And third, one may discover conceptual truths by empirical means—as was the case for 'whales are mammals'.

[169] We ignore, for ease of exposition, questions concerning the relation between genera like *whale* and species like *humpback* or *fin whale*—but see our brief remarks concerning this relation in §4.2 (esp. on p. 133).

[170] Looking at the example more closely, one might argue that the earlier judgment 'whales are fish' is, in a way, not mistaken at all, because the conception of 'fish' used merely indicated the manner of movement. For an interesting discussion of 'folk biology' that might support such an assessment see Atran [1998].

By way of comparison, notice that the notion of epistemic analyticity does not have these implications. We cannot be wrong about which are the epistemically analytic truths precisely because, so the thought goes, understanding what the terms occurring in them mean entails knowledge of the relevant truths. They depend on our conceptions, i.e., on *our grasp* of the concepts in question.[171], It is impossible, therefore, to discover new epistemically analytic truths concerning a given conception by way of empirical research: either it confirms the standing conception, in which case the epistemically analytic truths remain the same, or it leads to additions to or changes in the standing conception, in which case the corresponding epistemically analytic truths change as well.[172]

Turning now to aprioricity, it is clear that similar observations apply. For, on our view, many of the things we find out empirically are in fact conceptual truths—it is just that we need empirical investigation in order to adequately grasp the right concepts. Nevertheless, conceptual truths are, in a very straightforward sense, knowable *a priori*: you need only fully grasp the concept in question to know the corresponding truths. All the same, e.g., that *there are* whales is not knowable *a priori*, because knowing *what* whales are does not tell you *that* whales are. What is potentially confusing here is that there is, again, an epistemic version of the distinction between the *a priori* and the *a posteriori* as well, which is not based on conceptual truth but rather on (potentially imperfect) conceptions. On this second reading, what is knowable *a priori* is just what can be known on the basis of one's current conceptions, without further empirical investigation. Again, this notion is unstable, because graspings of relevant concepts may differ such that for the one person, further empirical investigation is necessary to arrive at a particular conceptual truth while for another this is not the case (or, at least, such a scenario is not ruled out).

Finally, we should note that the notion of *a posteriori* necessity, or of *de re* necessity, which played such an important role in the revival of metaphysics since they were introduced by Kripke [1972] and Putnam [1975], simply *is* our notion of conceptual necessity. The label '*de re* necessity' is more apt than '*a posteriori* necessity'—*de re* necessities about something are truths whose truth is guaranteed by the fundamental

[171] Alternatively, one might construe a private notion of analyticity which rests not on the public meaning of the terms involved but rather on the private conceptions involved. Then, one cannot be wrong about such analytic truths because they follow from one's own conceptions. Of course, one may question whether we have infallible access to our own conceptions, but that is irrelevant to our discussion.

[172] One might, of course, empirically investigate people's conceptions of something in order to find out about them—that would presumably be part of sociology or psychology or perhaps linguistics. The so-called 'experimental philosophers' are engaging in this kind of (unphilosophical) investigation. See, e.g., Knobe and Nichols [2008, 2013].

concept under which that thing (*res*) falls (amongst other things), whereas the *a posterioricity* they are said to enjoy amounts, to repeat, to a potentially unstable feature that depends on the epistemic situation of particular people. Of course, insofar as the label '*a posterioricity*' was just meant to flag a departure from the empiricist conflation of the necessary, the *a priori*, and the analytic, we wholeheartedly agree. On our interpretation that departure is a move away from the Modern picture and towards the Aristotelian picture: concepts are no longer entirely separated from reality but rather form part of it ('meanings ain't in the head').

In the previous chapter (§2.1.1), we appreciatively mentioned Kit Fine's view that metaphysical necessity is grounded in the essences of things. That is, a thing's essence guarantees the truth of certain statements about it. This idea meshes nicely with the view we have developed here: a thing's essence, on our view, *is*, at least in part, the concept under which it fundamentally falls (in the sense developed in the previous chapter, see §2.2.1), and it is the contribution that this concept makes to the statements in question which helps guarantee their truth. Essential truths are, then, conceptual truths, and therefore metaphysical truths are conceptual truths as well.

It seems plausible, once we have come thus far, to simply identify metaphysical and conceptual modality. But there are reasons for resisting this thought. For instance, one could argue that there are lots of *conceptual* truths about unicorns, but it seems strange to hold that these are *metaphysical* truths as well (or, perhaps, they are metaphysical truths in a qualified sense to be found in a proper theory of fiction). In other words, it may well be that metaphysical modality is best conceived of as a restriction on conceptual modality—in this case, a restriction to those concepts that could really be instantiated, in some sense or other. Here as in the case of logical modality we do not commit ourselves to a specific view on what the restriction might be—'could really be instantiated' is thus deliberately vague.

3.4 Concepts and Essences

What emerges from our discussions is the following picture. The fundamental idea we have discovered to underlie notions like logical form and logical consequence is that of truth guaranteed by content alone. We call those truths that are true in virtue of their contents, on an unrestricted (and non-epistemic) reading, *conceptual truths*. The corresponding notion of *conceptual consequence*, then, is a relation that holds between a set of premises and a conclusion just if the truth of the conclusion is guaranteed by the content involved conditional on the assumption that all premises

are (plainly) true. If that relation holds, we say that the argument in question is *conceptually valid*. The guarantee that is involved in all of these notions is a form of modality: we call it *conceptual modality*.

Given these core notions of conceptual truth and conceptual consequence, we can go on to differentiate: there are those conceptual truths and validities whose truth and validity is guaranteed by the content involved even if you leave out several aspects of that content. We can isolate truth-functional aspects of propositional connectives, and leave out all further content, to study their logic, or we can isolate aspects of predication and quantification to study their logic. What is left of the content of statements if you leave out all but the contributions to content such logics concentrate on may be called their logical form. The notion of logical form is, then, relative to such a restriction on the content that is taken into account.

This perspective on the logical properties motivates a fresh look at isolating kinds of content, with an eye on conceptual interrelations and distinctions that logical orthodoxy omits. We thus propose to look beyond the traditional boundaries of what counts as purely logical content and isolate aspects of content that are important from a metaphysical point of view. We offered some reasons for thinking that such metaphysically interesting varieties of truth-guaranteeing content are indeed left out if we stick to the traditional idea that predicate logic is the paradigm of logic—even within the realm of typical predicate-logical notions such as 'predicate', 'thing', and 'predication'. We could insist on calling the kinds of conceptual study we have in mind 'logic', and thus talk about, e.g., a logic of causality and temporality, a logic of life, etc. However, that is a merely terminological question, and since our departure from what the academic discipline of logic is taken to consist in is quite radical, it seems better not to hijack the term for our purposes—we are already pushing the envelope with our non-standard use of the term 'conceptual'.

A critical reader might object to our conclusions in this chapter by arguing as follows: we have simply stated that conceptual truth, truth guaranteed by content, is the fundamental notion, without having explained in any way how this truth-guaranteeing is supposed to work. Well, on this count we plead guilty as charged. Our aim has been precisely to work towards the insight that no account of this truth-guaranteeing can be given without circularity. In fact, we suspect that the Tarskian approach to logical validity is so widely embraced precisely because it seems to provide the kind of understanding or account that our critic demands: the guarantee that undergirds logical truths, so the Tarskian, is just their being true in every model. Ironically, the very same notion that makes our imaginary critic raise her eyebrows

is here used without hesitation both in establishing the appropriate space of models and in attempting to show that the account succeeds, as we have seen. Put differently, our aim has been to develop something like an act of 'conceptual ostension or, if you like, of eidetic intuition' [M. Thompson 2008, p. 105]: just look what happens as you think through the various concepts we attempted to develop, and then ask yourself again what exactly it is you are asking for when you insist on the sort of account our critic requires.[173]

On a more positive note, we may compensate for our supposedly failed analysis by pointing towards the practice of logic: apparently, we do quite well at identifying interesting aspects of content to study in separate logical systems, using various approaches. All such studies increase our grasp on how the truth-guaranteeing works in those specific cases. We hope to add to this grasp on some metaphysically interesting scores in the second part of this essay, albeit not (or perhaps not yet) with the formal rigor that we have come to associate with the discipline of logic. We will distinguish, on a fairly general level, between three different clusters of concepts, each with its own typical form of predication together with accompanying construals of objects and their properties. Let us proceed to our conceptual-metaphysical study of abstract entities, of concrete objects, and of living beings.

* * *

[173] An interesting argument for the 'indispensability' of the very idea of thoughts being reasons for inferring further thoughts—a close cousin of our notion of conceptual consequence—can be found in Stroud [2011, ch. 4].

Part II

Conceptual Realism: Metaphysics

Chapter 4

Conceptual Gears

What we have argued for in the first part of this essay has been of an exceedingly general nature. We have argued for an understanding of metaphysics based on a rejection of the currently orthodox Modern picture in favor of our alternative, Aristotelian picture. We moved on to elucidate and defend the central tenet of this Aristotelian picture—that essences are concepts—by defending a suitable version of essentialism and by developing a suitable understanding of conceptual truth. We have refrained, however, from making any claims as to what exactly the metaphysical implications of this Aristotelian picture are.

In this second part, we will gradually descend from these generic heights to somewhat more concrete levels. The path downwards is dictated by our results so far: given that essences are concepts, a differentiated study of the conceptual provides insight into the workings of essence. In what follows, we thus distinguish and study several different ways of thinking, and thence arrive at a more substantial and differentiated sketch of a metaphysical view based on the assumptions of the Aristotelian picture than what we have been in a position to formulate up until now. Our main source of inspiration for this second part of our project is Michael Thompson's work, in particular his idea of 'conceptual gears'.[174]

Thompson chooses this term, 'gear', because it captures in a pictorial way the phenomenon he has in mind: when shifting conceptual gears, we move from one cluster of concepts to a similar cluster of concepts, every element of which has undergone

[174] See M. Thompson [2008, esp. p. 56]. His own views on the topic will become relevant in §4.4 below and in chapter 7, especially §7.1. We have found further inspiration for our views in work on classifications of the sciences published around 1900—in particular, Steiner [1886, 1894] and Ostwald [1908, 1929].

a certain metamorphosis and differentiation. The idea is that the richer conceptual structures which are to be found in higher gears enjoy a certain independence, in the sense that they do not reduce to the more parsimonious or perspicuous conceptual structures we find in lower gears—where reduction is understood along the lines sketched in §2.2.1. This irreducibility shows up in a certain kind of circularity: we can explain (or 'define') a concept belonging to one specific gear only by invoking other concepts from within that same gear.

This chapter introduces three such conceptual gears in some detail. As we have argued in chapter 3, especially in §3.3.2, it is possible to distinguish between various kinds of contributions made to the overall content of a statement or thought not just by the concepts involved but also by the style of predication involved, by the way in which something is said of something. To each of our three gears there correspond different such styles of predication, along with different sorts of predicates and different sorts of things which one can predicate of. These form the focus in our introduction of each of the gears below.

In this chapter, we will provide little argument as to why we should accept these three gears as parts of our considered, metaphysical world view. Arguments to that effect will be provided as we move on to discuss the bearing that our division into three gears has on contemporary metaphysical debates in the ensuing three chapters, which are devoted to a less expository, more engaged in-depth discussion of the mentioned three conceptual gears. There, we argue that many contemporary discussions in metaphysics and in the philosophy of science can be understood as clashes between thinkers who feel at home within different conceptual gears. This insight enables, *inter alia*, a better understanding of how metaphysical realism, wedded as it is to the Modern picture, works: by elevating one particular way of thinking, one particular conceptual gear (or even just one aspect thereof), to the special status of representing reality as it is in itself, the realist metaphysician comes into conflict with those who are more in favor of a different way of thinking, of a different conceptual gear. On the other hand, the study of such conflicts, which will be our main business in the three chapters to follow, also greatly enhances our grasp of the different conceptual gears, and hence helps us sketch in rough outline the metaphysical position to which the Aristotelian picture gives rise—a position which does not thus privilege one particular conceptual gear over the others, but is rather built on the insight that the diversity of conceptual gears just is the diversity of reality.[175]

[175] Earlier, we mentioned Aristotle's recognition of the need for diversity in the way we think about

Interaction with contemporary debate is thus to be found in the ensuing chapters, not in the present one. After a somewhat more elaborate introduction of the notion of a conceptual gear in §4.1 below, we turn to the task of providing a perspicuous presentation of our three gears in §4.2, §4.3 and §4.4, respectively.

4.1 Conceptual Gears

We distinguish three essentially different ways of thinking about reality, three distinct clusters of concepts or conceptual gears. This amounts, on the Aristotelian picture, to distinguishing three aspects of reality. These three conceptual gears range from the more abstract and general to the more concrete and specific. Roughly, the first conceptual gear concerns *entities* in the most general sense, mathematical entities being the typical examples. The second conceptual gear concerns concrete, physical *objects* existing in space and persisting through time, and the third concerns living *organisms* and their specific unity.

Our claim is that each of these three conceptual gears, each of these levels of thought, consists in a distinct, irreducible cluster of concepts, marked out by typical conceptual relations and, correspondingly, typical forms of predication, which give rise to characteristic valid argument forms.

Interestingly, there are families of metaphysical views within the contemporary (realist) metaphysical debate that stay entirely within the first or entirely within the second gear, and attempt to apply it to all of reality, to the exclusion of higher gears—resulting in reductive views on the relevant aspects of reality. In particular, broadly Humean metaphysical views, paradigmatically David Lewis's, stick to the first conceptual gear.[176] Opposed to such views, we find varieties of so-called anti-Humeanism, such as David Armstrong's metaphysical program, or metaphysical views that center on powers or dispositions, which have rapidly gained influence over the last decades.[177] Anti-Humeanism is typically wedded to the second conceptual gear, and many of its adherents wish to stay entirely within that gear, resulting in

reality—see fn. 159 on p. 113.

[176] See, e.g., Lewis [1986a, 1991], Sider [2001, 2012], and Hawley [2002]. First-gear thinkers that do not fit into the Lewisian camp include, e.g., Chalmers [2012] and Williamson [2013]. In calling such views Humean, we do not mean to take a stand on what the actual Hume thought (although we will have occasion to briefly look at his views; see §5.2); we are merely following conventional ways of denoting a specific family of metaphysical views. Perhaps we do best to hereby stipulate that our 'Humean' is short for 'neo-Humean'.

[177] See, e.g., Armstrong [1983, 1997, 2010], and, for powers-based metaphysics, see, e.g., Martin [1997], Martin and Heil [1999], Heil [2003, 2005], Mumford [1994, 1998, 2004], Molnar [2003], Bird [2005, 2007], and Mumford and Anjum [2011].

reductive views on all third-gear aspects of reality. We will have occasion to mention and discuss various first-gear/Humean metaphysicians on the one hand, and various second-gear/anti-Humean metaphysicians on the other hand, in the course of our investigations into first- and second-gear thought in chapters 5 and 6. Chapter 7 deals with the third gear and those second-gear thinkers that wish to reduce it.

There are, of course, many more metaphysical views, that fall outside the categories of both first-gear/Humean and second-gear/anti-Humean metaphysics. Some of them, in particular neo-Aristotelian views, may not be very far removed from the kind of view we are developing in this essay—both in its understanding of metaphysics, as developed in Part I, and in its metaphysical implications, which we are about to develop in this second part.

It may seem, now, that we are arguing for a metaphysics based on these three gears *exclusively*. That, however, is not the case. It may well be that further investigations show that the entire gearbox contains more than just three gears, with extra gears either above or between the three to be discussed here. Indeed, we would encourage such further differentiation where it makes sense. But that is not the purpose of this essay: for our delineation of the metaphysical outlook to which the Aristotelian picture gives rise it is enough if we restrict ourselves to the mentioned three gears—that provides us with ample material, as will become clear.

Still, we have reasons for focusing on just these gears—our choice is not arbitrary. We already mentioned that comparing the first and the second conceptual gears provides materials for understanding what exactly is at issue in the controversies between Humeans and anti-Humeans, which make up an important part of the contemporary metaphysical discussion, not only within metaphysics proper but also within the philosophy of science. And we include the third conceptual gear because it illustrates the limitations of both Humeanism and anti-Humeanism, leading to a rejection of the currently popular reductionist and physicalist inclinations. It is here that the peculiarity of the metaphysical outlook based on the Aristotelian picture will be most visible. And, it is here that we expect our results to be at their most controversial, but also most promising: recognition of the third conceptual gear as a distinct and irreducible conceptual nexus may help us understand our own place as human beings living in a physical world. It is to this effect that Thompson embarked on his conceptual/metaphysical investigations: his work extends from the realm of the living to that of intentional action and further to that of ethical practices.[178] We leave out these further gears for obvious reasons of scope.

[178] Such is, in fact, the structure of Thompson's [2008] book into three parts.

That said, let us now start with a first rough scheme of our three conceptual gears:

	First gear	**Second gear**	**Third gear**
form of predication:	tenseless	tensed/aspected	normative
applies to:	entity	object	organism
predicates involved:	feature	state/process	life-process
sortals involved:	type	natural kind	life-form
mode of being:	to exist	to persist	to live
engaged in:	—	process	life-process
characteristic:	geometry	causality/temporality	teleology/normativity
	Humeanism	*anti-Humeanism*	

The notions used will be explained as we go along—this scheme is meant only to give a preliminary idea of what is to come.

Before we move on, however, the following three remarks are in order. First, the various gears are not mutually exclusive; in fact, if higher gears apply, so do lower ones. E.g., a living organism is also a persisting object and an existing entity.

Second, through various clever philosophical tools of paraphrase and analysis, those espousing only first- or second-gear thinking try their best to make it seem as if they can say everything which a defender of the second or third conceptual gear wants to say as well. We have seen some of this in §1.1.1; we will see more of it in due course, and, once we have surveyed all the gears and recognized their importance for our understanding of reality, it will become clear that such strategies are typical of the Modern picture, but in fact entirely unmotivated once we have abandoned that picture in favor of the Aristotelian picture.

And third, it may take some effort to unearth the conceptual gear that one or another philosopher is in fact using. One needs to look through the words used, as it were, which, by our second remark, can be misleading in this respect, in order to discern the actual way of thinking that is being employed. In fact, this applies also to our own exposition of the diverse gears below: we invite the reader to look through the terminology we use, the appropriateness of which one may quarrel about, in order to discern the ways of thinking, the forms of thought, we intend to bring to light.

4.2 The First Conceptual Gear

As we said, the first conceptual gear is especially suited for mathematical entities, like numbers, geometrical figures, and sets, but also graphs, algebras, structures, logics,

etc. We reserve the word 'entity' to denote first-gear entities; 'object' or 'physical object' will play a similar role in the second gear. Likewise, we use 'type' for sortal concepts in the first gear; 'kind' or 'natural kind' will play a similar role in the second conceptual gear. Nothing in particular hinges on our choice of terms; we merely need a reasonable terminology to explicate the differences between the various gears.

The first thing to note, with respect to the abstract entities of mathematics, is that they are completely static. They don't change, they have their properties and stand in relations atemporally. Moreover, they do not have any properties merely possibly, nor could they lack properties they in fact have (try imagining the number 17 not being a prime number). There is no distinction between contingent and necessary features. We may thus conclude that, because temporal and modal distinctions do not apply, these entities are *atemporal* and *amodal*.

One might think that these bold claims are simply false. Easy counterexamples, like the property 'x is Jim's favorite prime number', may be ignored—these are simply not at home within the first conceptual gear since they involve concepts from higher gears (i.e., the concept of a person and of an attitude towards some entity). But there are more difficult cases as well. In model theory, for instance, we can define different models with the same domain such that one and the same entity from that domain has different properties in each of those models. Isn't that a case of contingency? Well, what we have here is in fact a way to *simulate* contingency within a first-gear setting (see §5.1, p. 164 for more on simulation). But on the face of it, it is not 'real' contingency: the models are individuated by their constituent domains and interpretation functions, hence there is nothing contingent about a given model ascribing certain properties but not others to a given entity from its domain. And there is nothing contingent about a given entity being ascribed different properties by different models either.

Despite this amodality, there still is a distinction between what is *essentially* true of any first-gear entity and what is not. Given the results of chapter 2, we may trust that the abstract entities typical for first-gear thought have their defining features essentially, while also having features that are not essentially had. Adapting Fine's [1994b] example to a purely first-gear one, we may say that it is essential to the set $\{3\}$ that it contains the number 3, but not essential to the number 3 that it is contained in $\{3\}$. Thus, what *type* of entity one is dealing with, the sortal under which it ultimately falls, still carries essentialist weight in accordance with the essentialist inference.

Our example points to a conjecture concerning the first conceptual gear that we should formulate right at the start. Consider the thought that $3 \in \{3\}$: it is an essential

truth with regard to {3} because it follows from the type of entity {3} is—a set. Hence it is a conceptual truth. Every truth that is essential to something or other is a conceptual truth, as we have seen in chapters 2 and 3. The conjecture now is that every first-gear truth is in fact a conceptual truth. This conjecture arises because we have said that there is no contingency in the first gear, hence nothing which just happens to be the case. The thought is thus that the truth of every first-gear truth is guaranteed by the concepts involved (such is our definition of conceptual truth; see §3.3.1). We call it a 'conjecture' because we do not wish to commit ourselves to the thesis that the conjecture in fact bears out—we leave that to those concerned with the study of first-gear entitites (of 'abstract objects', as it is usually called). Therefore, in what follows we will talk as if there is room for truths that are not conceptual truths in the first gear. If our conjecture turns out true, this will have no interesting effects on the first-gear forms of thought we are about to identify.

Let us specify, for concreteness, a simple but illustrative toy example for the rest of this section: triangles within the geometry of the Euclidean plane. We will pretend 'triangle' to be an ontologically fundamental type of entity, for purposes of illustration.[179] Within the Euclidean plane, a *polygon* can be defined to be a closed figure consisting of line segments connected end-to-end. A *triangle* can then be defined to be a three-sided polygon (where we may or may not allow for limiting cases such as sides of zero length). Alternatively, one may define a triangle to be a polygon with three vertices (where we may or may not allow for limiting cases such as vertices of zero or 180 degrees). Of course, having three vertices follows from having three sides and *vice versa*, so one might wonder which of these definitions is 'more fundamental'. Though this is an interesting question for those developing an ontology of abstract objects, we will not delve into it here but merely note that simply taking the conjunction of such alternative equivalent definitions will do for present purposes, and also yields the comforting result that triangles are both three-sided and three-angled by their very nature.

We immediately touch upon questions that involve our conjecture that there are only conceptual truths in the first gear. That conjecture makes sense of the fact that so many conceptual interrelations and so many interdefinabilities have been established within and between the most diverse areas of mathematics. We thus deem it very likely that the essences of abstract entities are far more encompassing than we usually think, leaving only little room (if at all) for non-essential features, hence also for truths

[179] What determines fundamentality, in the first conceptual gear, is an interesting question, which falls, however, outside of the scope of the present essay.

that are not conceptual truths. This would explain the air of conventionality that is involved in providing definitions of mathematical entities (e.g., in terms of either three-sidedness or three-angledness): one needs to grasp only a very small part of the essence of a mathematical entity in order to derive the rest of its essential features, hence *which* part we start with is a matter of choice.[180]

Given its atemporality and amodality, the sorts of predication involved in the first conceptual gear are fairly simple and undifferentiated, but they nevertheless go beyond the Fregean style of predication as mere function application, which is familiar to us from its central role in first-order predicate logic. To be more precise, we distinguish two basic sorts of predication, which we briefly introduce below. For simplicity's sake, we only discuss the basic case of monadic predication; we ignore relations and other complicating factors for now. It should also be kept in mind that our aim is not to construct a logic; we aim to capture forms of thought typical for the first conceptual gear in an informal way.[181]

We call the first and most familiar sort of predication *feature predication*: '*a* is *F*' simply says that some entity, *a*, has a certain feature, *F*—where the 'is' is to be read tenselessly. An example: 'Triangle *t* is right-angled'. This sort of predication coincides with the Fregean way of understanding predication, and is the only sort that plays a role in first-order predicate logic. The second, less familiar case is that of *sortal predication*: '*a* is a φ', which says that *a* belongs to the fundamental type φ. If true, such a thought *always* amounts to a conceptual truth, as we argued in chapters 2 and 3, which means that the concept expressed by φ is the (general) essence of *a*. An example: '*t* is a triangle'.

Insofar as there is room for truths that are not conceptual truths, these will thus have to involve feature predication (which is not to say that *no* truthful feature predication is conceptually true, of course). Our above-mentioned conjecture is supported by the fact that it is difficult to find examples of such 'non-conceptual' truths. Examples of conceptual truths involving feature predication, such as '*t* is three-sided' and '*t* has three vertices', are much easier to find. Whether or not right-angledness also belongs to triangles as a matter of conceptual necessity depends on how they are individuated—again, our purpose is not to defend any particular view on geometrical entities, just to survey the typical forms of thought involved.

[180] Mathematically speaking, it is interesting to find out which parts of an essence entail which other parts—hence the observation that mathematicians are interested in many different proofs for the same result.

[181] Formalization is an interesting project in its own right, as is illustrated by, e.g., the work of E.J. Lowe, who develops a 'formalized sortal language' FSL in his [2009, chs. 11–12]. However, Lowe does not differentiate between different conceptual gears as we do.

Now, both of our types of predication can be applied not just to entities but also to types, yielding thoughts of the forms 'φ's are F' and 'φ's are ψ's', respectively. Examples include 'triangles are three-sided' and 'triangles are polygons'. In the case of feature predication applied to a type, we may again distinguish between those truths of that form which are conceptual truths—in which case being F is included in the account of φ's—and those truths of that form that are true for other reasons. An example of the latter might be 'triangles are set-members'. Sortal predication applied to a type, on the other hand, always yields a conceptual truth when true.

A few clarificatory points are in order with regard to sortal predication. We claim that sortal predication, when true, always yields conceptual truths, while in the case of feature predication things might turn out one way or the other. But what about putative sortal predication involving non-fundamental sortals? Again, different cases have to be distinguished. The first and easiest case is that involving too general sortals (genera, as Aristotle would say), such as 't is a polygon'. These are conceptual truths for quite straightforward reasons—the fundamental sortal under which the thing in question falls will include the more general sortal (i.e., triangles are polygons). The second, opposite case is sortal predication involving sortals that are reducible to fundamental sortals (in our sense, see §2.2.1), such as 't is an isosceles triangle'. In such cases, we go beyond merely stating what the entity in question most fundamentally is (an isosceles triangle is a triangle *having two equal sides*), and hence should be understood as involving feature predication (perhaps as a conjunction of sortal predication and feature predication, to yield 't is a triangle and t has two equal sides'). As in the case of ordinary feature predication, such thoughts may amount, when true, to conceptual truths, but need not. The third and last case to be distinguished is sortal predication involving putative sortals that cut across various fundamental sortals. Consider, for instance, the sortal *insider*, which we hereby define to apply to any geometrical entity lying within the boundaries of a given triangle, such that 't is an insider' applies to triangles, circles, points, squares, etc. Now, obviously, this is just feature predication in disguise—such 'sortals' tell us nothing about what a given thing fundamentally is, only something about the features it has.

Considerations such as these in fact strongly suggest that it is wise to restrict the term 'sortal' to fundamental sortals such as *triangle* and the general sortals they include, such as *polygon*—when it is one's business to uncover fundamental forms of thought, categories derived from linguistics (such as 'sortal') are only helpful up to a certain point. We hereby adopt this suggestion: 'sortal' will henceforth only name

sortal terms expressing fundamental types of entities and the genera under which they fall (in the higher gears, this will transform into natural kinds and life-forms respectively—see below).

Having thus briefly clarified the two basic sorts of predication involved in the first conceptual gear, it is interesting to note that, from a standard predicate-logical perspective, sortal predication is simply treated as feature predication, while both sortal and feature predication applied to types is treated as universally quantified feature predication. Thus, 'triangles are polygons' is regimented to yield 'For all x, if x is a triangle then x is a polygon'. As we observed in our brief discussion of the shortcomings of predicate-logical form in §3.3.2, first-order predicate logic leaves out interesting aspects not only of *what* is predicated and of *of* what is predicated, but also of the style of predication involved. When operating within the first gear, this assimilation of different forms of thought to just feature predication is largely innocent, although it still obscures interesting aspects of first-gear thought (in particular, the essentialism to which it gives rise by way of the essentialist inference). When shifting to higher gears, however, it becomes more and more clear that the distinctions we have been exploring here amount to ever more marked differences.

The assimilation of distinct sorts of predication to just feature predication, which is typical of predicate logic, may be extended, e.g., by including a set of times into the domain and making all predication relative to such times by adding an argument place for times to each predicate. Thus we arrive at a first-gear simulation of temporality, which is in fact a second-gear notion—in the second conceptual gear, as we will see below, temporality further differentiates the various sorts of predication.[182]

To illustrate the interaction between the two types of predication and their application to either entities or types, let us look at two very simple examples of valid arguments which make use of all of them. The validity we are interested in here is validity on the basis of our forms of predication, in abstraction from the entities, features, and types involved.

Here is our first example ('G1' stands for 'gear 1'):

	P_1	Triangles are$_{\text{sortal}}$ polygons	φ's are$_{\text{sortal}}$ ψ's
[Arg. G1.1]	P_2	t is$_{\text{sortal}}$ a triangle	t is$_{\text{sortal}}$ a φ
	$\therefore$	t is$_{\text{sortal}}$ a polygon	t is$_{\text{sortal}}$ a ψ

This argument (which we refer to as argument [G1.1]) operates entirely on the

[182] Here we witness the origin of the contemporary metaphysical dispute between eternalism (first gear) and presentism (second gear). See the next section below, and in particular §6.2 for detailed discussion.

level of sortal predication. All of the thoughts involved are conceptual truths. On higher gears, this type of reasoning turns out to remain largely the same. Notice that if we do not restrict ourselves to validity in virtue of types of predication, that is, if we consider conceptual validity in the broad sense of §3.4, no premises are needed to reach the conclusion—the conclusion is itself a conceptual truth, after all.[183]

Now, let us move on to our second example:

	P_1	Triangles are$_{\text{feature}}$ three-sided	φ's are$_{\text{feature}}$ F
[Arg. G1.2]	P_2	t is$_{\text{sortal}}$ a triangle	t is$_{\text{sortal}}$ a φ
	$\therefore$	t is$_{\text{feature}}$ three-sided	t is$_{\text{feature}}$ F

We mention this particular example because its counterparts in higher gears turn out to be importantly different—feature predication, as opposed to sortal predication, turns out to undergo considerable differentiation upon shifting to the second and third conceptual gears. We take the first premise, P_1, to be a conceptual truth deriving from the concept of a triangle: it involves feature predication. The second premise, P_2, states that the the object t falls under the (fundamental) sortal 'triangle', hence it is a case of sortal predication. From these two, it follows that t is three-sided: again a case of feature predication, in this case itself amounting to a conceptual truth. Note that an instance of this type of argument which features a first premise that is *not* a conceptual truth (if such there be) will result in a conclusion that is not a conceptual truth either. As before, the argument is valid if we restrict ourselves to the types of predication involved; if we consider conceptual validity in the broad sense the conclusion needs no premises to be arrived at since it is a conceptual truth.

Let us now put the basic forms of thought we distinguished into a scheme:

The First Conceptual Gear

	Applied to entities	*Applied to types*	
Sortal predication	a is a φ	φ's are ψ's	(all conceptual truths)
Feature predication	a is F	φ's are F	(some conceptual truths)

[183] Perhaps a brief clarificatory remark is in order here. One might think that there is no way to read off from the *name* 't' that it is a name for a triangle, and that therefore the conclusion itself is not a conceptual truth. We, however, do not mean to be operating on such a purely syntactical level here. On our view, as we sketched it in chapter 3, the thought expressed in the conclusion contains the essence of t, which is the fundamental sortal concept under which t falls: 'triangle'. *That* thought is a conceptual truth.

What we have ignored in our discussion is not just relational predication but also predication applied to features (or relations), such as 'Three-sidedness is a shape property' or 'Being two meters apart is a distance relation'. Here we could again distinguish between sortal predication (involved in these examples) and feature predication (for which it is difficult to find examples within the first conceptual gear, but perhaps 'there are 5 Platonic solids' will do). We have ignored these further aspects of first-gear thought because they do not really help us grasp what is distinctive about the diverse conceptual gears, and also for reasons of scope.

When thinking within this first conceptual gear, one can achieve a level of precision and rigor that is typical for mathematical thought. This feature may count as one of the major motivations for attempts to understand all of reality from within this first gear. It is interesting, in this respect, to note the fact that analytic philosophy grew out of a distinctively mathematical nest, so to speak.[184] Symptomatic for this tendency is the skeptical attitude towards all things modal and temporal (in the more robust sense to be explained below) as well as the conviction that first-order predicate logic is 'the' logical system, the universal framework for talking about anything at all.[185] On reflection, one may find that this tendency makes a lot of sense when one is doing mathematics, hence also when one is applying mathematics to problems in other sciences—but not when one considers the aims of those sciences themselves, given that they aim at understanding parts or aspects of our temporal-modal world.

4.3 The Second Conceptual Gear

When we shift to the second conceptual gear, the notion of an entity transforms into that of a concrete, physical *object* that exists in space and persists through time.[186] It is obvious, then, that the atemporality which was distinctive of the first conceptual gear now vanishes—or rather, that the second conceptual gear can be characterized as involving temporality. And, along with temporality, contingency now enters the scene too: lots of truths concerning concrete, physical objects are not necessarily

[184] As Millgram writes: 'There's nothing to get philosophers to jump on a bandwagon like a formal representation with mathematical cachet' (see Millgram [2009, p. 5]).

[185] These symptoms call to mind, of course, the work of W.V. Quine. It is no accident that the most influential advocates of first-gear metaphysics in this sense have been strongly influenced by Quine—in particular, David Lewis.

[186] Such physical objects are sometimes named 'continuants' or 'substances' [see, e.g., Lowe 2009, p. 18f and 89ff, Wiggins 2001, Preamble, §3]. These notions, however, tend to encompass not just phyiscal objects in our sense but also living organisms, which we turn to in our discussion of the third conceptual gear below. Therefore, we prefer 'physical object' for the second-gear notion we have in mind.

true. For instance, that two physical objects—say, your desk and your chair—are now standing next to each other in precisely *this* spatial arrangement is, although probably explainable in causal terms (on which more below), surely not a necessary truth in the sense in which it is necessary that $3 \in \{3\}$.

Temporality, as we intend it here, is a characteristic of the second conceptual gear. Hence we should take great care not to let our first-gear habits get in the way of a proper grasp of this notion. Those first-gear habits include that of conceiving the temporal dimension as being relevantly similar to the spatial dimensions—which we might call the geometrization of time.[187] No such first-gear simulation can be found in the second conceptual gear. Instead, the temporality involved here shows up in styles of predication peculiar to the second gear. For a start, feature predication needs to be qualified temporally now, e.g., 'the ball *is* red', 'the ball *was* red', 'the ball *will be* red'. Those are all systematically interrelated but different ways of combining some feature—redness—with some object—the ball—in thought.[188]

This example, however, makes use of a kind of property, 'being red', that is still close to first-gear types of properties. It is what we call a *state-property*, and is in that respect like 'being a prime number' or 'containing 3 as a member'. But the temporality involved in the second conceptual gear also introduces an entirely different kind of property, which is not to be found in the first conceptual gear: that of producing or undergoing change. E.g., 'the ball is rolling'. We call such properties *process-properties*. And with these process-properties come *processes*: a new type of entity. Processes are the activities of physical objects, and, conversely, physical objects are entities capable of engaging in various activities.[189]

In this section, we adopt as our stock example of a physical object 'a ball'. The example is a bit awkward because balls are artifacts and in that respect involved in more than just the second conceptual gear—artifacts are products of humans and hence are designed for a specific purpose. On the other hand, the example is also an illustrative one: balls are interactive objects with which we are all familiar. That

[187] We will be more specific on this score in the next two chapters. §5.1 sketches first-gear simulations of temporality, while §6.2 contains an in-depth investigation into what a truly second-gear understanding of time amounts to.

[188] Temporal or tense logic can be conceived of as the study of how exactly these tenses interrelate. It has famously been inaugurated by Arthur Prior [see, e.g., his 1967]. As with most formal investigations, there is some pressure towards a purely first-gear understanding of tense, even amongst those who 'take tense seriously' by treating it as a modal operator instead of assimilating it to location in space. Examples can be found among advocates of the so-called 'thin red line' [see, e.g., Øhrstrøm 2009, Malpass and Wawer 2012]. We will critically discuss Prior's attempt at capturing time as a second-gear notion in §6.2.4.

[189] There are, of course, defenders of 'pure process' metaphysics, such as the followers of Whitehead [1920, 1929]. That would be a metaphysics restricted to only part of this second gear. See, e.g., Sellars [1981a,b,c], Rescher [1996, 2000], and Bickhard [2011a,b].

is why we stick to it anyway and ignore its artifactual aspects. If the example turns out to be too confusing after all, we advise to substitute something else—perhaps a planet, or a star, or a mountain. We should add that the second conceptual gear includes a category apart from physical objects, namely, stuffs—which we will ignore, however, for reasons of clarity and scope.[190]

It is convenient to distinguish feature predication involving state-properties from feature predication involving process-properties. We will call the first type of predication *state-predication*, and the second *process-predication*. Separating these two as distinct types of predication makes sense because process-predication allows for a further kind of differentiation which is absent from mere state-predication. Although we do not wish to rely on purely linguistic observations for our conceptual investigations, it is helpful to characterize the differentiation we have in mind with the help of the distinction in English (and other languages) between perfective and progressive *aspect*. Together with the temporal differentiation we already mentioned, we thus have six forms of process-predication:

	Progressive	*Perfective*
Present	The ball is rolling	The ball has rolled
Past	The ball was rolling	The ball had rolled[191]
Future	The ball will be rolling	The ball will have rolled

The conceptual differentiation we have in mind here, and which seems to be nicely expressible by way of this linguistic distinction, is that between being engaged in a certain process on the one hand, and having completed a certain process on the other hand. These two different ways of saying something—rolling—of something—the ball—differ in the inferences they allow for. To say that the ball has rolled is to say that there *has been* a process with a beginning *and* an end, so that it is no longer there. To say that the ball is rolling is to say that there *is* a process with a beginning which is still ongoing and hence does not (yet) have an end.[192] We will further illustrate these inferential differences below.

As it might be, our ball stops rolling because it hits another ball which then starts

[190] But see Laycock [2006] for an interesting discussion of the formal peculiarities of statements about stuffs as opposed to statements about objects and statements about pluralities.

[191] According to Rödl [2012/2005, ch. 5, §2], this particular form does not belong on our list: 'the ball had rolled' does not unify 'the ball' with 'rolling' differently from 'the ball has rolled'. The only difference lies in that by using 'had' instead of 'has' one can go on saying something from the point of view of some past moment—and that is a pragmatic difference, not a conceptual or metaphysical one. We are inclined to agree, but leave the naive scheme as is because discussion of the issue would take us too far afield.

[192] This relatively simple distinction turns out to become rather important in the third conceptual gear, as we will see in the next section.

rolling. Surely, then, it was *possible* for the ball to continue rolling—if only the other ball had not been there to stop it, it would have rolled on. Here, we discover another important ingredient of the second conceptual gear: a species of modality. We call the kind of modality involved here *causal modality*. It is best illustrated by invoking another typical second-gear notion that has been lurking in the background of our exposition so far: that of a *power*, or disposition—or, to use the classical Aristotelian term, that of a *potentiality*.[193] Our ball has the power or potentiality to roll: it is 'rollable', we might say. Whenever its power to roll manifests itself, a process of rolling starts. What is required for its power to roll to manifest itself in some actual episode of rolling may be a rather complicated story—but a proper strike with a cue when placed on a billiard table will most probably do the trick. But then, of course, the manifestation might be extremely short-lived, e.g., when some evil philosophers have set up an H-bomb to go off just when the tip of the cue touches the ball, instantly vaporizing it just as it starts to roll. . . .

The point of this short story is to bring out the peculiarity of the causal modality that is involved here: an object's powers tell us what kinds of processes it can engage in, and hence implies a range of in-principle causal *possibilities* for the object. We may call this aspect of causality, following Aristotle, *formal causation*.[194] This formal aspect of causation determines not what will happen but rather which processes are possible. Now, typically, powers only manifest themselves under certain conditions. E.g., a ball which is at rest needs to be pushed or pulled in some way, by some force, in order to manifest its power to roll.[195] And that is, again following Aristotle, efficient causation: it results in the realization of one of the possible processes determined by the formal aspect of causation. But even if such proper circumstances are in place, that act as an efficient cause, nothing follows as to *how far* the process in which the manifestation of a power consists will unfold: it is always in-principle possible for there to be some kind of interfering factor which prevents the process from reaching one or another stage of its development. This is the notorious *ceteris paribus* character of many claims about the causal activities of concrete physical objects.[196] It makes

[193] See, e.g., *Metaphysics* [Aristotle 1998], Book IX.

[194] The formal cause is one of the four causes Aristotle distinguishes—see *Physics* [Aristotle 1996], Book II.5, and *Metaphysics* [Aristotle 1998], Book V.2. However, our use of this label in no way aims to be faithful to Aristotle's views—it is, rather, the connection between the sortal concept under which a thing falls, on our view, and Aristotelian forms, that moves us to use this notion.

[195] We ignore the exceptional theoretical possibility of spontaneous rolling envisaged by Norton [2007]—it illustrates an indecisive or spontaneous power (see below).

[196] We briefly remark on a famous criticism of *ceteris paribus* laws, which holds that they are either false or trivially true—see, e.g., Lange [1993] and Earman and Roberts [1999]. On the one hand, so the thought goes, any attempt to spell out the *ceteris paribus* clause for a given law will lead to a falsehood because

the idea of causal necessity difficult to grasp: when put in the right conditions, our ball will manifest its power to roll with causal necessity, yet it is still possible that the ensuing causal process is interrupted, even already at the very point at which it was started, as our example involving an H-bomb makes clear. We suggest that the best way of understanding the necessity involved is to stop looking for a certain causal *result* and to focus on the *process* involved. It is the causal *process* of rolling that necessarily comes about when our ball is placed in the right kind of conditions, not its actually managing to roll for any given stretch, for we may always construct cases in which the H-bomb goes off just before that precise stretch is reached.[197]

It is not our aim here to engage in more detailed discussions of these matters, but it is important to mention one further complication: the second conceptual gear allows for 'spontaneous' powers which do not require any determinate manifestation conditions. A case in point is the power of certain heavier chemical elements for spontaneous fission. ^{240}Pu, for instance, has the nasty habit of spontaneously splitting sometimes, which may cause a nuclear weapon to prematurely detonate—hence the plutonium used in such weapons is to be as low in ^{240}Pu as possible. On the other hand, there are conditions which will trigger ^{240}Pu's fission power—which is what is made use of in the case of controlled detonation of nuclear weapons.

Insofar as there are such spontaneous powers, the future is open. To illustrate what this 'open future' involves, assume that at the present moment, several causal processes are unfolding. These processes might continue as they do, leading to changes in the features and overall layout of the physical objects, in turn leading to certain powers manifesting themselves, which means that new processes begin to unfold, some of which might interfere with or stop processes that were still unfolding, etc. As long as there are no spontaneous powers involved, the course of events is completely determined. But if a spontaneous power is involved, the determinacy is broken: if it manifests itself spontaneously, the course of events will be different from the way it would be if it did not manifest itself.

In fact, such indeterminacy does not even require spontaneous powers in the

the point of having the clause is that there are indefinitely many possible disturbances that cannot all be listed. On the other hand, leaving the clause unanalyzed comes down to stating that the respective law holds except where it doesn't—a triviality. We would argue that this skeptical dilemma rests on first-gear reductionism: it is assumed that laws take the form of universally quantified implications. Moreover, on our view, *ceteris paribus* laws are best understood to be about the outcomes of processes. They do not mention all the ways in which interference might occur, they only tell us how the process involved unfolds. That does not make them trivial: they still express the fact that the process in question *does* come about, however short-lived it may be.

[197] That is to say, a process of rolling does imply *some* stretch of actual rolling, but no determinate such stretch, however small it is. We come back to this issue below, see p. 146.

strong sense just illustrated: a power which is 'indecisive' in that it may result in a range of different processes in response to the same manifestation conditions also results in indeterminacy, and hence an open future, without being spontaneous in that sense. Determinism, as a second-gear thesis, is thus the thought that there are no such spontaneous or indecisive powers.[198] A spontaneous or indecisive power may allow for probability distributions over its possible outcomes—probability, in this sense, is thus a second-gear notion, even if it can be treated quite successfully in mathematical and hence first-gear terms.

We have arrived at a conceptual framework including two distinct, generic types of entities: physical objects on the one hand, and the causal processes they are engaged in on the other hand. Correspondingly, there are two types of *kinds*: kinds of physical objects, and kinds of process. The former includes, e.g., 'ball', 'electron'[199], 'star', etc., and the latter includes, e.g., 'rolling', 'repelling', 'shining'. Presumably, these two types of kinds are closely interconnected: it is a conceptual truth, for instance, that balls roll, that electrons repel each other, that stars shine (given our understanding of conceptual truth; see especially §3.3.3). Such generic conceptual truths can be read as indicating power-possession: balls have a power to roll, electrons have a power to repel, stars have a power to shine. Finally, we may say that laws of nature systematically describe which kinds of things have which powers, just when those powers are manifested, and how exactly the ensuing causal processes interact. The generic conceptual truths that indicate power-possession are thus laws that govern the formal aspect of causation, while the laws that govern the efficient aspect of causation take forms like 'In circumstances C, process p will be initiated'. As we said earlier, such laws will typically leave open whether or not the process will indeed unfold or be interfered with (recall the philosophers with their H-bomb). Such is the *ceteris paribus* character of much second-gear thought.[200]

We mentioned conceptual truths of the form 'balls roll' or 'balls can roll': conceptual truths that indicate power-possession by instances of a certain kind of object. Balls have a power to roll irrespective of the circumstances they are in: they have an in-principle power to roll. This means that there is no role for tense to play in

[198] We take a closer look at the second-gear notion of an open future in §6.2.6.

[199] Electrons are problematic, because it is unclear whether or not they are physical objects at all. Like most microphysical entities, they have mighty strange features and seem to occupy a domain somewhere between individuals and stuffs, if that is conceivable at all.

[200] E.J. Lowe defends a theory of laws of nature on which they are taken to express generic ascriptions of power possession in our sense. See, e.g., Lowe [2006, ch. 8, esp. §8.6, 2009, ch. 9]. (Lowe remarks that a law of nature is, 'quite literally, . . . a law concerning the *nature* of things of the kind in question'—see [2006, p. 127].)

such generic thoughts: 'balls can roll', in this sense, is a tenseless thought (and a conceptual truth). Correspondingly, there is a tenseless power-ascribing form of thought that applies to individual objects instead of to kinds—for instance, 'this ball can roll'. That is, of course, different from the thought 'this ball is rolling': the former tenselessly ascribes a power to the ball, the latter ascribes tensed manifestation of that power. Now, one might think that there surely are circumstances in which a given ball can*not* roll: for instance, when it is hanging in empty space far removed from any other macrophysical object. That is true, but it involves a way of ascribing powers that differs from the tenseless, in-principle power-ascriptions we have just seen. We may call it *situated* power-ascription. Situated power-ascription is tensed: if we say 'this ball can roll' in this situated sense, we mean that the circumstances it is in *right now* allow for manifestation of its in-principle capacity to roll. Tenseless power ascription corresponds to in-principle possibility, while situated power-ascription corresponds to what we may call *real possibility*.[201] This notion of real possibility comes in degrees that depend on the allowed temporal scope: it is a situated, real possibility for me, for instance, to learn Swedish, even though I cannot do so within the hour; it is not in the same sense a real possibility for me to *speak* Swedish—but it *is* a real possibility for me to do so if we widen the temporal scope such that I can first learn and then speak Swedish. For any language that is learnable, it is really possible for me to learn it to an acceptable degree, but it is not really possible for me to learn *all* learnable languages—not even if we widen the temporal scope to its maximum.[202] Real possibility depends, then, on in-principle possibility: speaking loosely, it is the intersection of the ranges of in-principle possibilities of all physical objects involved.

Corresponding to the tenseless, in-principle power ascription we may identify a tenseless form of state predication. Like tenseless power-ascriptions, such truths can be derived from generic conceptual truths: balls are round, hence our ball is round in this tenseless sense. It is also round in the tensed sense, of course.[203]

Let us put together our sketch of second-gear forms of thought by way of a

[201] A nice exposition of this notion of real possibility, including a good formal rendering, is being developed by Rumberg and Müller [2013] and Rumberg [n.d.].

[202] On both the Ockhamist and the Peircean views on future contingents, as described by Prior [1967, ch. VII], there is no graduality in the notion of real possibility—but on both views the degrees *can* be introduced by restricting the temporal scope within which witnesses for the future claims need to be found on the relevant history (or on all histories). Again, see Rumberg and Müller [2013] and Rumberg [n.d.].

[203] Balls can be broken or flattened, in which case they will no longer be round. That, however, has to do with their artifactual nature: artifacts often have a certain intended shape, hence if that shape is altered, the artifact is malfunctioning ('broken' in a normative sense). We ignore this complication here; as we said earlier we mean to conceive of balls as merely natural objects, such that breaking or flattening them simply means destroying them.

scheme similar to the one we presented for the first conceptual gear in the previous section. The scheme is, of course, more complicated. Instead of feature predication we now have both state predication, which itself may be differentiated temporally, and process predication, which may also be differentiated by aspect. In the scheme, we indicate tensed predication, encompassing past, present and future tense, with a superscript $^+$, while absence of this $^+$ indicates tenseless predication. As said, the scheme encompasses two types of entities, physical objects (a) and processes (p), which are interrelated by the conceptual truth that processes are the activities of physical objects. We can apply sortal predication as well as state predication to instances of both of these categories, while process predication forms the connecting form of thought. Physical objects and processes fall under sortal concepts which are natural kinds of object (φ) and of process (V, for 'verb') respectively, and both physical objects and processes can be said to have certain (qualitative) features (F). They are connected by the various forms of process predication: tenseless in-principle power ascription, tensed situated power ascription, and tebsed ascription of an actual activity, which also takes aspect. As before, we ignore relational predication etc. for simplicity's sake. For reasons of clarity, we divide our scheme of the second conceptual gear into two halves, one for the various forms of predication as applied to objects and processes, and one for the various forms of predication as applied to kinds of objects and kinds of processes. Here is the first half:

The Second Conceptual Gear

— Applied to objects/processes —					
	tensed/aspected	*tensed/aspectless*	*tenseless*		
Sortal				a is a φ	Objects
State		a is$^+$ F	a is F		
Process	a is$^+$ V-ing a has$^+$ V-ed	a can$^+$ V	a can V		
State		p is$^+$ F	p is F		Processes
Sortal				p is a V-ing	
	Some are conceptual truths			*All conceptual*	

The upper part of this scheme concerns different types of predication as applied to objects, which are mirrored in the lower part for processes. The middle line shows the forms of thought that connect these two types of entities. We haven't listed all the modifications that state predication and process predication can take due to temporality (past, present, and future tense)—these are indicated with a superscript $^+$, as said (see the little scheme we provided on p. 138). As we argued in the previous section, sortal predication always yields conceptual truths when true, while the other types of predication may or may not express conceptual truths.

Now moving on to the second half of the scheme:

The Second Conceptual Gear

— Applied to kinds of object/process —			
Sortal		φ's are ψ's	Objects
State	φ's are F		
Process	φ's (can) V		
State	V-ings are F		Processes
Sortal		V-ings are W-ings	
	Some conceptual	*All conceptual*	

The organization of this half of the scheme is similar to that of the first half, except that all the tensed and aspected forms of thought are missing here; these differentiations only show up when predication is applied to individual objects or processes. It is clear, thus, that this second half of the scheme involves a less radical change from the first conceptual gear than the first half. Relatedly, the various interesting forms of bringing together object and process that we found on the level of individual predication reduce to just one in this second part of our scheme.

On the basis of this more or less complete outline of second-gear forms of thought, we can have a look at a few corresponding argument forms. In our discussion of the first conceptual gear above, we presented two argument forms, [G1.1] and [G1.2] (see p. 134), where the first involved sortal predication and the second involved feature predication. With regard to sortal predication, the situation in the second conceptual

gear is largely like it was in the first gear. But with respect to feature predication things are different now. Hence, we survey some of the different argument forms to which the differentiation of feature predication into state-predication on the one hand and process-predication on the other hand gives rise. Our first example is an easy one involving state-predication:

	P_1	Balls $\text{are}_{\text{state}}$ spherical	φ's $\text{are}_{\text{state}}$ F
[Arg. G2.1]	P_2	b $\text{is}_{\text{sortal}}$ a ball	b $\text{is}_{\text{sortal}}$ a φ
	$\therefore$	b $\text{is}^{(+)}_{\text{state}}$ spherical	b is_{state} F

When dealing with state-predication, we see that things remain largely as they were in the first conceptual gear, except that the conclusion may now be read both in its tenseless and in its tensed version (hence the $^{(+)}$). The argument is a fairly traditional example of a (logically valid) syllogism (but as we will see, its third-gear version turns out not to be valid).

An analogous argument involving process-predication is more interesting:

	P_1	Balls $_{\text{process}}$ roll	φ's $_{\text{process}}$ V
[Arg. G2.2]	P_2	b $\text{is}_{\text{sortal}}$ a ball	b $\text{is}_{\text{sortal}}$ a φ
	$\therefore$	b $\text{can}_{\text{process}}$ roll	b $\text{can}_{\text{process}}$ V

What makes this case different from state-predication is that we have only one way of applying process-predication to sortals, as in P_1, but three ways of applying that same style of predication to objects (apart from the further differentiation due to tense and aspect): in-principle power ascription which is tenseless, situated power ascription which is tensed, and power manifestation ascription which is also tensed. Hence from one and the same generic power-ascribing premise P_1, combined with the sortal-ascribing premise P_2, three different individual power-ascribing conclusions may be thought to follow. Yet it is valid for only one of these conclusions: the in-principle power ascription employed in [G2.2]. For compare the corresponding invalid varieties of this argument:

	P_1	Balls $_{\text{process}}$ roll	φ's $_{\text{process}}$ V
[Arg. G2.3]	P_2	b $\text{is}_{\text{sortal}}$ a ball	b $\text{is}_{\text{sortal}}$ a φ
	$\not\therefore$	b $\text{is}^{+}_{\text{process}}$ rolling	b $\text{is}^{+}_{\text{process}}$ V-ing

	P_1	Balls $_{\text{process}}$ roll	φ's $_{\text{process}}$ V
[Arg. G2.4]	P_2	b $\text{is}_{\text{sortal}}$ a ball	b $\text{is}_{\text{sortal}}$ a φ
	$\not\therefore$	b $\text{can}^{+}_{\text{process}}$ roll	b $\text{can}^{+}_{\text{process}}$ V

Obviously, a ball is not rolling just because it is a ball. Nor need it be the case that

a ball is situated such as to be able to roll just because it is a ball—as we said, only the in-principle, tenseless conclusion that the ball can roll is licensed.[204]

To illustrate the workings of the second conceptual gear, for instance with regard to the interaction of tense and aspect, many more examples could be presented. We restrict ourselves to just three examples here. Here is the first one:

[Arg. G2.5]

P	b has$_{\text{process}}$ rolled	b has$_{\text{process}}$ V-ed
$\therefore$	b was$_{\text{process}}$ rolling	b was$_{\text{process}}$ V-ing

When a process of rolling has come to an end, there was a process of rolling going on at some earlier point in time (earlier, that is, than the point in time at which premise P is true): this interdependence between perfective and progressive aspect is fairly straightforward.

Things are slightly less clear when considering the second example:

[Arg. G2.6]

P	b is$_{\text{process}}$ rolling	b is$_{\text{process}}$ V-ing
$\therefore$	b has$_{\text{process}}$ rolled	b has$_{\text{process}}$ V-ed

If a process of rolling is under way, it will be true to say that the ball has rolled. The most natural situation to think of here is, of course, one in which the ball *stopped* rolling such that we can say 'the ball was rolling, and then it stopped, so it has rolled'. But even if the ball is still rolling on, such that there is no end point (yet), it will be true to say that the ball has rolled for a certain stretch. That is because even if the process of rolling is interfered with immediately after it started, there will always be a small interval in time between the onset of the process and the occurrence of the interfering factor. If there is no such interval, that is, if the interfering factor *coincides* with the (putative) onset of the process, there is no process, for then the circumstances for the manifestation of the corresponding power are not fulfilled, in which case no process comes into being at all.

As an aside, note that we do not accept *as basic* second-gear processes such as 'rolling down to the river'. For such processes, our last argument seems to fail: a ball may of course be rolling down to the river without thereby having rolled down to the river. We understand such cases not as concerning a process different from rolling as such, but rather as including additional *de facto* information about the situation. As we will see below, failure of the above argument illustrates the third conceptual gear, where we find a different sort of process that does include its own proper end-point, not in a merely *de facto* manner but in a normative way (see §4.4, p. 154).

[204] But notice that the argument from 'stars shine' via '*s* is a star' to '*s* shines' does seem valid even on a tensed reading of the conclusion. That, however, has to do with the nature of stars and their shining activity, and hence is a validity that is based on more than just the forms of predication involved.

The third example goes beyond the simple forms of thought we have surveyed, but also illustrates a slightly more complex interaction between various forms of predication, in particular, a transition from generic, tenseless truths to a tensed conclusion:

[Arg. G2.7]

P_1	Balls $_{\text{process}}$ roll when in circumstances C	φ's $_{\text{process}}$ V in C
P_2	b is$_{\text{sortal}}$ a ball	b is$_{\text{sortal}}$ a φ
P_3	b is$^{+}_{\text{state}}$ in circumstances C	b is$^{+}_{\text{state}}$ in C
$\therefore$	b is$^{+}_{\text{process}}$ rolling	b is$^{+}_{\text{process}}$ V-ing

This argument does lead to a conclusion involving actual process-ascription because it involves both a specification of the circumstances in which the process is indeed initiated and a (tensed!) premise stating that these circumstances are right now in place. Arguments such as this one presumably play a role in scientific experimentation, where the aim is to determine which of various candidates for P_1 is true. To repeat, we take thoughts like P_1 to express laws of nature.

Now that we have a decent grasp of the second conceptual gear, we can have a brief look at the difference between first-gear thinking and second-gear thinking. A helpful way to do so is by considering Laplace's famous demon. As the story goes, the Laplacean demon, who knows the full layout and distribution of physical objects and their properties throughout reality at the present moment, including the laws of nature, is in a position to calculate how reality will unfold in the future as well as how it came to be the way it is.[205] This sounds like a paradigmatically second-gear story (on the assumption that the world is deterministic, that is, that there are no spontaneous or indecisive powers): we have physical objects endowed with certain deterministic powers governed by the laws of nature, and we have, on the basis of a complete description of the present state of the universe, both an absolutely accurate prediction of how things will develop in the future and an absolutely accurate reconstruction of how things were in the past.

However, Laplace himself was engaged in the (highly successful) project of 'mathematicizing physics', which we may characterize as a move towards applying first-gear thought to second-gear phenomena. Hence, Laplace's story can also be read in a first-gear way, roughly as follows. Imagine a huge, three-dimensional array of dots, together with a set of (deterministic) equations telling how to construct an ordered series of further such huge, three-dimensional arrays on the basis of the given one. We may now call the given three-dimensional array the 'present moment', the set of equations the 'laws of nature', and we may glue all the three-dimensional arrays we

[205] See Laplace [1995/1814, pp. 2–3]. Laplace does not talk of a 'demon' but rather of an 'intelligence'.

can construct on its basis together into one big four-dimensional array. We may call the extra dimension thus introduced 'time', and the full four-dimensional picture the 'entire history of the world'. Understood in this way, we have a purely mathematical, first-gear story, which is only dressed with second-gear terminology—or, to be more precise, the story provides simulations of second-gear notions within a first-gear setting.

This illustrates not only the difference between first- and second-gear thinking, but also the ease with which we may slide from the one to the other. We do not mean to suggest that there is something wrong with using first-gear thinking to approach issues for which the second conceptual gear is appropriate to use: recall that where higher gears apply, so do lower gears. Indeed, isolating those aspects of second-gear phenomena to which first-gear thinking can be sensibly applied has proven to be very helpful in the development of physics and other sciences. However, things are different when such applications of first-gear thinking are thought to provide a way to *reduce* second-gear phenomena to first-gear phenomena. In the following two chapters we will discuss this issue at length.

4.4 The Third Conceptual Gear

We have seen a considerable increase in conceptual richness in moving from the first to the second conceptual gear. Moving on now to the third conceptual gear, the change turns out to be comparatively subtle, but still comprehensive. In fact, what changes can be summarized by use of just one concept, which is characteristic of the third conceptual gear. It is the concept of what we will call *teleology* or *natural normativity*. This characteristic, which is again a species of modality, is added to the temporal and causal modalities we found in the second conceptual gear. We start by considering what effect the addition of teleology has on processes.

Apart from having merely causal processes, which simply unfold as they do unless they are interfered with, we now find *life-processes* which *are supposed to* unfold in a certain way, hence something is *wrong* if interference occurs. Biology textbooks are full of examples illustrating this type of normativity. Let us have a look at one example for concreteness.

The life-process of photosynthesis occurs, among other places, within the mesophyll inside plant leafs. The cells of that tissue contain lots of chloroplasts, which in turn contain two types of systems where photosynthesis takes place: photosystem I and photosystem II (PsI and PsII). These systems guide incoming photons to their

respective reaction centers. Within the reaction center of a PsII-system, we find a specialized chlorophyll-a molecule sensitive to light of wavelengths up to 680nm. Such a molecule is called a P680 molecule. Now, as *CliffsQuickReview of Plant Biology* explains:

> The energy absorbed by a P680 molecule of chlorophyll in the reaction center of PsII converts the molecule to its excited state and raises the energy level of an electron, which moves to an outer orbit and is lost from the molecule. An acceptor molecule in the adjacent electron transport system—a chain of alternately oxidized and reduced compounds—accepts it, and immediately moves it down the transport chain, losing some energy in each transfer. The P680 molecule, having lost an electron, is unstable; it returns to its stable ground state by drawing an electron from a water molecule using an as yet not clearly understood mechanism of photolysis. The P680 is then ready to accept another photon, and the process is repeated over and over. [Rand 2001, p. 100]

This quotation describes, without going into much detail, only one early stage of the full photosynthesis life-process, but that is enough to illustrate our point. The description looks in some respects like a purely chemical (i.e., second-gear) description of a certain chain of subsequent causal processes. But in other respects it does not. Most importantly, it is not a description of a *particular* chain of causal processes which just happen to subsequently occur, but a *generic* description of a unified life-process which occurs again and again in many different sorts of organisms. In the description, it is simply assumed that there is an 'adjacent electron transport system' present, and it is simply assumed that there is water available for the P680 molecule to extract an electron from. That is because the chain of causal processes described plays a certain role for the organism: the P680 molecule, for instance, is *meant* to capture photons, although this is, of course, just one of its indefinitely many causal possibilities. At no stage does the life-process go beyond what is causally possible, of course, but the causal possibilities are exploited in such a way as to further the benefit of the host organism. The life-process of photosynthesis plays a certain role in the existence—the *life*—of the organism. It is *no accident* that the relevant further molecules are around inside the host organism; if one of them fails to be around, something has gone wrong, for then the process may very well no longer be of use for the plant in question. It will still be a process of photosynthesis, but it will be a failed such life-process. No failure in this sense is present in the second conceptual gear—indeed, it is not even *thinkable* in the second conceptual gear. Thus, whereas in the second gear we discovered a type of modality, causal modality, that is peculiar in that it manifests what we called a 'ceteris paribus' character (see p. 139 above), here we discover a type of modality, teleological or normative modality, that is peculiar

in that what it necessitates can nevertheless fail.

Another way to illustrate the typical unity and teleology of life-processes is to point out, as Thompson does, that it always makes sense to ask 'and what happens next?' [see M. Thompson 2008, p. 41]. In our example, we could ask what happens at the other end of the electron transport system. The answer will be that the life-process continues *as it should*, namely such as to eventually store the captured light energy in chemical bonds, yielding sugar and starch. Such a question will receive a radically different answer when we have ordinary causal processes in mind. Suppose that a bunch of P680 molecules are just floating around on the surface of the sea, capturing photons—what happens next, in such a case, depends simply on what happens to be around. There is nothing in the physical photon-capturing process itself that tells us one way or the other. When dealing with a life-process, on the other hand, there *is* something in the life-process itself that points with a certain normative force beyond its present point of development—in particular, it locates the life-process itself within the broader context of the life of an organism (or of an ecosystem, for that matter).

This brief introduction of the difference between second- and third-gear thinking will immediately give rise to lots of questions, many of which have to do with doubts about the need for a separate conceptual gear at this point—isn't what is *really* happening satisfactorily describable in purely second-gear (i.e., in physico-chemical) terms? For instance, it is observed that the sciences have informed us about lots of inorganic mechanisms that can be found in nature, e.g., the geological mechanism of mountain formation, or the meteorological mechanism that results in lightning, or even purely inorganic mechanisms of self-organization [see, e.g., Okamoto *et al.* 2007]. What makes these mechanisms different from the mechanism of photosynthesis, apart from a mere increase in complexity? For now, we can only mention that such inorganic mechanisms do not involve the kind of normativity that is found in the third conceptual gear: their stability and unity have a mere *de facto* character. As with our description of the Laplacean story at the end of the previous section, we may notice that application of second-gear mechanistic thinking to third-gear phenomena is not only possible but has helped enormously in terms of scientific progress—yet that is something different from taking a *reductive* stance towards third-gear phenomena in general. We will discuss these and related questions in much greater length in chapter 7; here we will continue our task of sketching the forms of thought associated with the various gears in rough outline.

Life-processes occur in the context of the lives of organisms. In the third conceptual gear, the notion of an *organism* thus takes the place of the notion of a physical

object, just as life-processes take the place of processes. Since organisms are also physical objects, they exist in space and persist through time just like every other physical object does. One can think about organisms using second-gear forms of thought, that is. Persistence through time, however, also transforms into something new in the third conceptual gear: to persist, for organisms, is *to live*.[206] Accordingly, the sortals that ultimately characterize living beings are likewise not the same as the sortals under which non-living physical objects ultimately fall—natural kinds—but instead something different: *life-forms*. The life-form of an organism includes what kind of life such an organism lives, how it develops from birth to death, what kinds of life-processes it engages in on every level of its existence, and also how it relates to other life-forms, its normal environment, etc.

Life-forms thus give rise to numerous conceptual truths, just as second-gear natural kinds and first-gear types do. And these conceptual truths, like those concerning second-gear natural kinds, comprise what we have called the formal aspect of causation above (see p. 139): life-forms determine what is and what is not possible for their instances. The normativity that is typical of the third conceptual gear, however, makes such conceptual truths have a rather different status. The difference we have in mind is best brought out by continuing the series of examples of valid arguments we gave while presenting the first and the second conceptual gear. We begin by considering a third-gear analogue of the example involving state-predication we gave in the previous section (see argument [G2.1] on p. 145 above). Whereas this type of argument remained largely the same as we moved from the first to the second gear, the natural normativity involved in the third conceptual gear does make a difference (we use 'should' to indicate that normativity):

	P_1	The horse is_{state} four-legged	φ's $\text{are}_{\text{state}}$ F
[Arg. G3.1]	P_2	h $\text{is}_{\text{sortal}}$ a horse	h $\text{is}_{\text{sortal}}$ a φ
	$\therefore$	h should be_{state} four-legged	h should be_{state} F

There are, of course, horses that have lost one leg, hence the conclusion that horse h is four-legged cannot be drawn in its simple, non-normative form—only in its specifically third-gear, normative form: horses that have three legs *should have* four legs, hence they are defective in missing one leg. We see that the normativity here differentiates between two distinct types of thought involving state predication: the

[206] And, given that to exist, for physical objects, is to persist, we may as well say, with Aristotle, that to exist, for a living thing, is to live: 'for in everything the essence is identical with the cause of its being, and here, in the case of living things, their being is to live, and of their being and their living the soul in them is the cause or source.' See *De Anima* [Aristotle 1984, p. 27], II.5, 415b12–14.

factive one that *h is* four-legged, and the normative one that *h should be* four-legged.[207]

Next, we consider a third-gear analogue, involving life-process-predication (indicated with subscript 'lp'), of the examples involving process-predication we gave in the previous section (see argument [G2.2] on p. 145 above, and its invalid variants [G2.3] and [G2.4]). Here we observe a normative shift of a different kind, which replaces the shift to 'in-principle' predication we noticed in the second conceptual gear:

	P_1	Sugar maples grow$_{lp}$ five-fingered leaves	φ's $_{lp}$ L
[Arg. G3.2]	P_2	m is$_{sortal}$ a sugar maple	m is$_{sortal}$ a φ
	$\therefore$	m grows/should grow$_{lp}$ five-fingered leaves	m (should)$_{lp}$ L

As in the second-gear analogue [G2.2], the conclusion is a tenseless one. But it isn't a mere in-principle possibility ascription, which might or might not get actualized, as is the case in the second conceptual gear: to say that sugar maples grow five-fingered leaves is to say that they *should* grow five-fingered leaves—the latter formulation merely explicates the normativity that is already present in the former one. Growing five-fingered leaves is a life-process belonging to what it is to be a sugar maple. Compare the process of burning: our sugar maple m of course can burn to the ground, but that *is* a mere in-principle possibility. Its burning would be a second-gear causal process, not a third-gear life-process—connecting the process of burning with a sugar maple in thought yields a second-gear thought, not a third-gear one, and hence will not be characterized by the normativity we are now focusing on.[208]

Of course, the conclusion is not the tensed one that the sugar maple is *in fact* growing five-fingered leaves—e.g., it might just not be the right time of year for growing leaves, or it might suffer from some kind of disease preventing it from doing so. It might even suffer from that disease as long as it lives, such that it is clear from the very start that it will *never* produce any five-fingered leaves: the conclusion, *in its normativized form*, will still be true of it.[209]

[207] We should stress that the fact/norm distinction we seem to be introducing here is not a real distinction. The kind of normativity we are interested in is one that does not *contrast* with facts, but rather *characterizes* certain facts—the third-gear ones, to be precise. To illustrate: it is perfectly fine to say, with a tenseless use of 'is', that this (three-legged) horse *is* a four-legged creature. That 'factive' thought is identical to the explicitly normative one that figures as the conclusion in our argument.

[208] Note, though, that fire can play an important role for certain life-forms, as it does, e.g., for giant sequoias and redwoods.

[209] As we observed in our presentation of the second conceptual gear, there is a tensed version of this type of thought as well, stating that in the obtaining circumstances, it is really possible that the object in question engages in the activity mentioned. A similar observation applies to the third conceptual gear, again with a normative modification: in the right circumstances, it *should* be possible for our sugar maple, say, to unfold its leaves; and if everything is as it should be for sugar maples, it *is* possible for our sugar maple to unfold its leaves.

Now let us move on to a somewhat more elaborate example involving life-process-predication, a third-gear analogue of the argument from the presence of appropriate conditions to the actual manifestation of a power (see argument [G2.7] on p. 147 above). Interestingly, a similarly normativized conclusion figures in this example as well:

[Arg. G3.3]

P_1	Sugar maples flower$_{lp}$ in early spring	φ's $_{lp}$ L in C
P_2	m is$_{sortal}$ a sugar maple	m is$_{sortal}$ a φ
P_3	It is^{+} early spring	m is$^{+}_{state}$ in C
$\therefore$	m should be$^{+}_{lp}$ flowering	m should be$^{+}_{lp}$ L-ing

Here, as in the second-gear case, the conclusion is a tensed one. But even in suitable conditions (i.e., in early spring), the process in question need not be invariably initiated, as in the case of second-gear processes. Our sugar maple may fail to flower because, e.g., it is ill (hence is not as it should be), or because it has been fooled by being placed in an artificially controlled environment. Now, why is this failure of process-initiation not simply due to the conditions C being inadequately specified? If we explicitly exclude diseases and artificial environments etc., would not the process of flowering come about with causal necessity, revealing its second-gear nature after all (see the relevant discussion in the previous section, p. 141)? Consider: balls don't care, so to speak, whether or not they will ever roll (ignoring, as we have done so far, the complications that arise if we take into account their status as artifacts), whereas sugar maples *do* care whether or not they will ever flower: they *should* flower. That the proper circumstances for flowering arise is part of being a sugar maple, while this is not true for balls and their power to roll. That sugar maples grow where they grow is, in this sense, no accident. Hence the only circumstances that are relevant to mention in P_1 are those that specify the proper episode in a sugar maple's life for the ascribed life-process to come about—i.e., early spring. Again, it is typical of third-gear concepts that they bring with them their own embedding within a wider biological context—which is, ultimately, the earth's biosphere in its entirety considered as a living 'organism'.

The second-gear argument [G2.6] (on p. 146 above), moving from a progressive premise to a perfective conclusion, is an interesting one in that it is valid in the second conceptual gear but not in the third, as our next example shows:

[Arg. G3.4]

P	m is$_{lp}$ growing five-fingered leaves	b is$_{lp}$ L-ing
$\not\downarrow$	b will have$_{lp}$ grown five-fingered leaves	b will have$_{lp}$ L-ed

As we observed, the problem is that life-processes themselves point towards their

own continuation. The life-process that our sugar maple *m* is engaged in itself is directed at its product: five-fingered foliage. Now, even if disease or other interfering factors prevent such foliage from actually coming about, the life-process itself is one of growing five-fingered leaves. It is just that the process does not unfold as it *should*, hence it may be true that at one point *m is growing* five-fingered leaves without this implying that at any point it will *have grown* five-fingered leaves.

The point deserves some further attention. One might think that similar things can be said of, say, a ball rolling down the hill towards the river: suppose that it is caught half-way such that it never reaches the river in the valley. Is it not true then that is *was rolling* to the river but never *has rolled* to the river? We will not deny that such might be a perfectly fine way in English to express our thoughts about cases like this. Language is, after all, a flexible medium for expressing our thoughts. The difference we are envisaging is best captured by stressing the teleology that attaches to life-processes but not to causal processes: the river just happens to be at the lowest point for the ball. If it were to roll down another hill, it would end in a different place. Not so for the sugar maple: if it were in different circumstances, it would (and should) still grow five-fingered leaves. Insofar as causal processes have an 'aim' or natural end-point, this will just be a *de facto*, contingent matter of fact. Life-processes, on the other hand, have intrinsic, non-contingent ends built into them.

Now that we have pointed out a few of the more salient features of the third conceptual gear, we may compose a scheme of forms of predication for this gear. In our examples above, we introduced a new predication-modifier, 'should', to indicate involvement of the normativity typical of the third conceptual gear. As with the shift from first- to second-gear forms of thought, the shift to this new, third-gear element has the most impact when considering the various forms of predication as applied to individual organisms. We now have not only state predication (F) and process-predication (V) but also life-process-predication (L). And, correspondingly, we not only have organisms and processes but also life-processes. Since explicating the diverse ways in which our sorts of predication apply to processes and to life-processes will only unnecessarily complicate the scheme, we leave them out, and only provide the forms of thought that involve the various sorts of predication as applied to individual organisms. For similar reasons, we also leave out a list of generic forms of third-gear thought, i.e., of forms of thought based on the various sorts of predication as applied to life-forms. As before, the superscript $^+$ indicates tensed predication, absence thereof indicates tenseless predication; to further simplify matters we now also add a subscript $_\times$ to encode aspect:

The Third Conceptual Gear

	— Applied to organisms —			
	tensed/aspected	*tensed/aspectless*	— *tenseless* —	
Sortal				*a* is a φ
State		*a* is$^+$ *F* *a* should be$^+$ *F*	*a* is *F*	
Process	*a* does$^+_\times$ *V*	*a* can$^+$ *V*	*a* can *V*	
Life-process	*a* does$^+_\times$ *L* *a* should$^+_\times$ do *L*	*a* can$^+$ *L* *a* should be able$^+$ to *L*		*a* (should) *L*
	Some are conceptual truths			*All conceptual*

The organization of the scheme is similar to that of the second conceptual gear (see p. 143). We should mention one peculiarity: all the normativized forms of thought, when true, amount to conceptual truths. This is because the norm is always provided by the life-form in question, and since life-forms are concepts, such normative truths are always true in virtue of the concepts involved.

4.5 Concluding remarks

We have surveyed three conceptual gears: that of purely abstract entities, that of concrete, physical objects, and that of living organisms. We have seen how the rather simple scheme of forms of thought we identified in the first conceptual gear evolves into a much richer such scheme in the second conceptual gear, and we noticed that this richness has everything to do with the characteristics of the second gear: temporality and causal modality. We found a further increase in richness when the forms of thought typical of the second conceptual gear undergo another transformation into the third conceptual gear, which is characterized by natural normativity or teleology. Thus we moved from first-gear existing entities and features and the types they fall under to second-gear persisting objects and processes, states and powers, and the natural kinds they fall under, to third-gear living organisms, life-processes, and the life-forms they fall under.

It has proven impossible to sketch all of these diverse forms of though involved

in the diverse gears without immediately running into a host of philosophically problematic issues. To some extent, we already engaged in brief discussions of these issues, albeit in a rather sketchy manner, even though we promised to postpone such discussions to later chapters. That was simply unavoidable. We should note that we will not be in a position to discuss *all* of the issues touched upon in this chapter in greater detail. For example, we will not engage in a discussion of the metaphysics of abstract objects, even though we did touch upon that topic in our discussion of first-gear thought. For us, first-gear thought is primarily of interest because it can also be applied to all of reality, to the exclusion of other forms of thought. In the next chapter, we will discuss such an approach to metaphysics, which we called *second-gear metaphysics*, and we will identify the typical metaphysical views to which this approach leads, such as four-dimensionalism, eternalism, and modal reductionism (see esp. §5.1). It turns out that there is a close connection between first-gear thought and Humean approaches to topics such as laws of nature (although this claim needs some qualification, which we provide in §5.2.4).

An exclusively second-gear approach to metaphysics, which only allows second-gear forms of thought yielding what we have called *second-gear metaphysics*, does not involve as big a transition as in the first-gear case, simply because the second-gear forms of thought, as we have introduced them, already apply to concrete objects. In chapter 6, we will thus be concerned to some extent with the growing dissatisfaction with Humean views (especially concerning causation; see §6.1), and also with the opposition to other typically first-gear metaphysical views (especially concerning time; see §6.2). We regard the contemporary debates that have thus arisen as indications that the second conceptual gear is being re-discovered after many decades of Humean hegemony within analytic philosophy.

Our discussion of the third conceptual gear in chapter 7 takes a different form: here it will be our aim to show that we should indeed endorse this conceptual gear in all its conceptual robustness, and resist reduction to second-gear simulations such as 'mere' mechanisms. Chapter 7 is thus the most controversial chapter of the second part of our essay, but also the most interesting one in terms of the overall project: we will see that it brings us in a position to develop a better understanding of how concepts can find expression in concrete phenomena, and thereby enables a better understanding of our central thesis that essences are concepts.

* * *

Chapter 5

The First Gear: Humeanism

In the previous chapter, we introduced the notion of a conceptual gear, and surveyed the typical forms of thought associated with the three conceptual gears we intend to explore within the second part of this essay. By way of orientation, let us repeat the overall scheme we provided at the beginning of the previous chapter here:

	First gear	**Second gear**	**Third gear**
form of predication:	tenseless	tensed/aspected	normative
applies to:	entity	object	organism
predicates involved:	feature	state/process	life-process
sortals involved:	type	natural kind	life-form
mode of being:	to exist	to persist	to live
engaged in:	—	process	life-process
characteristic:	geometry	causality/temporality	teleology/normativity
	Humeanism	*anti-Humeanism*	

This chapter is devoted to a study of metaphysical views that restrict themselves to first-gear forms of thought: first-gear metaphysics. The first section contains a general description of what such views look like by reference, largely, to the school of metaphysical thought inaugurated by David Lewis. We argue that many typical features of this Lewisian take on metaphysics illustrate the underlying refusal to allow anything other than first-gear forms of thought. In particular, we focus on how first-gear metaphysicians deal with second-gear phenomena. We divide our discussion into two parts, one focusing largely on temporality (§5.1.1), the other on modality (§5.1.2). Related topics, such as contingency, change, persistence, and mereology, are covered along the way.

The second section (§5.2) delves deeper into one interesting topic where the clash between first-gear thinking and second-gear thinking is especially vivid: laws of nature. We start out by considering Hume's famous reasons for being skeptical about 'necessary connections' (§5.2.1), in order then to move on to a more focused consideration of contemporary Humean views on laws of nature (§5.2.2). We uncover the first-gear character of such views by reviewing the metaphysical options for what we call the 'mosaic' that underlies laws of nature on a Humean view (§5.2.3). On reflection (§5.2.4), we come to see that our identification of first-gear metaphysics with Humeanism needs to be qualified: the Humean may endorse *at most* first-gear metaphysics, but is also under considerable pressure to reject metaphysical realism entirely in favor of some version of metaphysical anti-realism.

5.1 First-Gear Metaphysics

The first conceptual gear is typical in that it does not involve temporality or modality, as we said. Moreover, we noticed that the forms of thought one can find on the first gear are best suited for theorizing about abstract entities, paradigmatically mathematical entities. So in order to exclusively apply this conceptual gear to all of reality, which evidently (or at least apparently) does include contingency, temporality, and modality, the first-gear metaphysician has to come up with ways of introducing contingency, time, and modality into his picture without moving beyond his preferred conceptual framework. That is, he has to come up with ways of *simulating* such second-gear notions within a first-gear setting. It is instructive to see how these and related notions are accounted for within first-gear metaphysics, both for a better understanding of first-gear thought and for a better understanding of (the controversies over) the relevant metaphysical topics. In this section, we sketch, for illustration, the first-gear metaphysical view that has been systematically defended over many decades by David Lewis and his followers. Other versions of first-gear metaphysics are possible, of course—we will have occasion to illustrate the diversity of first-gear metaphysical views in our discussion of second-gear metaphysics in the next chapter.

5.1.1 Temporality

One popular way for first-gear metaphysicians to stick to their cherished first-gear forms of thought consists in adopting a construal of temporality containing three

elements: *eternalism*, *B-theory*, and *perdurantism*. We will dwell on all three of these in their opposition to presentism, A-theory, and endurantism at much greater length in our discussion of time as a second-gear concept in §6.2; for now, it will suffice to briefly indicate how these three doctrines result in a conception (or, rather, a simulation) of time that is acceptable by first-gear lights.[210] The most important ingredient in all three of them is the thought that times are locations in much the same sense as places are—which makes time suitable for a geometrical treatment. David Lewis opens his famous book *On the Plurality of Worlds* with a nice statement of the kind of view in question:

> The world we live in is a very inclusive thing. Every stick and every stone you have ever seen is part of it. And so are you and I. . . . There is nothing so far away from us as not to be part of our world. . . . *Likewise the world is inclusive in time*. . . . [N]o long-gone primordial clouds of plasma are too far in the past, nor are the dead dark stars too far in the future, to be part of this same world. . . . The way things are, at its most inclusive, means the way this entire world is. [Lewis 1986b, p. 1; emphasis added]

Moving beyond such a seemingly intuitive characterization, we may distinguish the three mentioned ingredients of the view Lewis (and many others) defends as follows. *Eternalism* is the view that time constitutes a dimension that is similar to the spatial dimensions: the spatial dimensions provide one with a multitude of spatial locations for things to be located at, and the temporal dimension extends that multitude. Space-time, so the thought goes, is one big manifold—the 'block universe', Lewis's 'very inclusive world'—within which everything past, present, and future is located. It is the container that contains all there is, in space and time alike. Eternalism is opposed to *presentism*, the thought that reality is not extended in time in this way, that the past and/or the future are not on an equal footing with the present moment, that they are in some sense not contained in the big container which is reality.

Related to this distinction, but not identical to it, is the opposition between B-theory and A-theory. *B-theory* is the view that temporal truths can be analyzed in tenseless terms: to say that something *is*, *was*, or *will be* the case is just to say that it tenselessly is simultaneous with, earlier than, or later than the moment of utterance (or evaluation). Thus, time is a collection of temporal locations, linearly ordered in terms of an earlier/later relation, at which such tenseless truths obtain. *A-theory*, on the other hand, is the view that the present is special in some way or other: temporal truths are tensed truths. The properties of being present, being past, and being future thus play an important role that is denied by the B-theorist (who insists on a tenseless

[210] For relevant references in addition to the few to be presented below, we also refer to §6.2.

description of reality). Given that A-theory insists on tensed predication while B-theory does not, it makes sense to infer that A-theory is a second-gear view while B-theory is a first-gear view (but we will qualify this statement on closer inspection; see §6.2.2). Notice that presentism thus requires A-theory, but not *vice versa*, while eternalism does not require B-theory—this will become important in our discussion of time in §6.2.

The most illustrative of the three first-gear doctrines concerning time, for present purposes, is *perdurantism*, which is intended to be an account of persistence, i.e., of how things manage to stay around through time. The term, though coined by Mark Johnston, was made famous by Lewis, who writes:

> Let us say that something *persists* iff, somehow or other, it exists at various times; this is the neutral word. Something *perdures* iff it persists by having different temporal parts, or stages, at different times, though no one part of it is wholly present at more than one time; whereas it *endures* iff it persists by being wholly present at more than one time. [Lewis 1986a, p. 202]

Perdurantists, then, say that objects divide up in time just like they divide up in space. Here, again, the idea that times are locations plays an important role: locations can be occupied, and how else can a thing occupy several locations than by having parts occupying each of them? Obviously, then, for Lewis, to 'exist at various times', his supposedly 'neutral' characterization of persistence, already implies a construal in terms of temporal *locations*, as is witnessed by his further characterization of endurance:

> Endurance corresponds to the way a universal, if there are such things, would be wholly present wherever and whenever it is instantiated. [ibid, p. 202]

Endurantists thus claim, in opposition to Lewis's view, that things are not temporally extended, do not have temporal parts. It is interesting to see how Lewis perceives this opposition to his preferred view, after having defended it for many years against many attacks:

> The only trouble with this hypothesis [i.e., perdurantism] is that very many philosophers reject it as counterintuitive, or revisionist, or downright crazy .... It is a mystery why. [Lewis 2002, p. 1]

It is quite easily seen how perdurantism becomes the only sensible view on persistence when times are thought of as locations to begin with, as Lewis does. The ordinary, common-sense thought on which the endurantist insists, viz., that it is literally the same thing that was first here and is now there (say), simply does not

make sense anymore: how can *the very same thing* occupy different locations? (We will argue for a version of endurantism that rejects the times-as-locations picture in §6.2.5.)

On a perdurantist view, objects persist through time by having temporal parts at various times (i.e., they perdure), just as they are spread out in space by having spatial parts at various places. And, just like no special form of predication is required for geometrical truths, so no special form of predication is required for temporal truths. Perdurantism thus implies B-theory: the tensed forms of predication are reduced to (tenseless) first-gear feature predication as applied to temporal parts. As Smart, another staunch defender of first-gear metaphysics, writes:

> It does not therefore seem to me that . . . there is any important difference between the tenseless 'is' of 'Eisenhower is President in 1956' and '7 + 5 is equal to 12'. Of course Eisenhower is a temporal entity, and so 'in 1956' has sense in relation to him, and numbers are non-temporal entities, and so there is no question of 'in 1956' in the case of the second proposition. This distinction can perfectly well be made explicit in the predicates of the two sentences and need not be done in the *copulae*. . . . A sentence of the form 'x is φ at t' is, of course, not timeless, any more than 'x is φ at such and such a place at t' is spaceless. Timelessness is not the same as tenselessness. '7 is a prime number' is both tenseless and timeless. (There is no sense in saying '7 is a prime number at t'.) [Smart 1963, p. 139]

Smart in effect argues that there is nothing essentially different about thinking of entities existing in time as opposed to thinking about abstract entities existing outside of time—and he does so by drawing an analogy with the spatial case: including a temporal location in something's having a property is just like including a spatial location. And that one can ascribe such locations only to physical objects, not to numbers, is no more surprising than that one cannot ascribe number-theoretic properties, such as primeness or evenness, to geometrical figures. The basis for all of this is the same: first-gear feature predication (and implicitly, we may add, sortal predication as well, in order to state to what type something belongs, e.g., *number*, *physical object*, etc.).

Returning now to perdurantism, we may note that it is quite naturally accompanied by another paradigm of first-gear thought, *classical extensional mereology*, as the correct theory of part and whole for physical objects: just as regions of space-time can be understood to be 'nothing over and above' their parts (their subregions down to their constituent space-time points), physical objects are 'nothing over and above' their parts too. The two cardinal claims of this mereological theory are, first, *universal composition*, according to which any multitude of given objects composes a further object, their mereological sum or fusion, and, second, *extensionality*, which

says that no two objects can have exactly the same parts (in other words, to any number of objects there corresponds at most one mereological sum). Here is Lewis, in his seemingly commonsensical mood:

> Given a prior commitment to cats, say, a commitment to cat-fusions is not a *further* commitment. The fusion is nothing over and above the cats that compose it. It just *is* them. They just *are* it. Take them together or take them separately, the cats are the same portion of Reality either way. [Lewis 1991, p. 81]

This mereological theory exemplifies the virtues of the first conceptual gear to an extraordinary extent: it is parsimonious, elegant, precise, and entirely atemporal and amodal. And, of course, formulable neatly within first-order predicate logic.[211]

The combination of perdurantism and extensional mereology not only provides for a conveniently geometrical simulation of persistence, but it also generates a conception of physical objects as the contents of regions of space-time, and hence enables an altogether geometrical understanding of physical objects *tout court*, again congenial to the first gear. We have already had occasion to discuss such an understanding of physical objects in chapter 2 (§2.3): it leads to spatiotemporal and mereological essentialism.

Defenders of this view on physical objects are often tempted to go even further in geometrizing their metaphysical view on physical objects by *identifying* the objects with the regions they occupy—so Sider:

> There is ... the question of whether ... spatiotemporal objects *occupy*, but are distinct from, regions of space-time, or whether they simply *are* regions of space-time. There is considerable pressure to give the latter answer, for otherwise we seem to gratuitously add a category of entities to our ontology. All the properties apparently had by an occupant of space-time can be understood as being instantiated by the region of space-time itself. The identification of spatiotemporal objects with the regions is just crying out to be made. [Sider 2001, p. 110]

Others, such as Quine [1964, 1976a], even go so far as to consider the option of identifying physical objects with n-tuples of reals, and thus with pure sets (recall our discussion of numberism in §1.1.1). It is typical for first-gear thinking that it tends to 'geometrize' or 'mathematicize' everything it reflects upon in this way.

Moving now to further aspects of this view, we may note that perdurantism allows for a conception of *change* that nicely fits the first conceptual gear: since persistence is taken to be perdurance, change is accounted for by noting that an object which

[211] We already mentioned classical extensional mereology before, in §2.3; see fn. 104 on p. 77. For a much more thorough discussion of extensional mereology and the role it plays for first-gear metaphysics, see Mulder [2010].

changes from being *F* to being non-*F* does so in virtue of having two distinct temporal parts, the first of which is *F* while the second one is not. Now, observe that change is, on this view, not different from spatial variation of properties: a poker which is *first* hot and *then* cold is very much like a poker which is hot *here* (on one end) and cold *there* (on the other)—both phenomena are explained by pointing to distinct parts of the poker, thus avoiding the uneasy conclusion that the poker itself, as a whole, has contradictory properties. This leads, however, to a natural objection to the perdurantist's conception of change: change is simply *not* analogous to spatial variation of properties, so the conception must be wrong.[212] We may again let Sider speak in reply to this objection:

> there is no reply making use of elaborate distinctions or theory. The [objection] may simply be met head-on. Change *is* analogous to spatial variation. Change *does* occur in virtue of unchanging facts about temporal parts. There are no good arguments to the contrary. [Sider 2001, p. 214]

What makes this objection-and-reply interesting, for our purposes, is that it clearly illustrates how the first and the second conceptual gear may clash. The objector comes to the discussion with a second-gear notion of change, while Sider operates entirely within the first gear and thus sees no way of construing change except as analogous to spatial variation of properties (viz. his remark: 'there are no good arguments to the contrary'). Again, the key assumption of the first-gear thinker is that times are just locations.

Notice that, by assimilating change to spatial variation, the perdurantist is forced to apply the same move to the very notion of a *process* (or *event*), of something happening (recall the important role processes plays in our sketch of the second conceptual gear in §4.3). And indeed, one of the most influential defenders of perdurantism, W.V. Quine, explicitly assimilates events to physical objects (in the perdurantist sense):

> A physical object, in the broad sense in which I have long used the term, is the material content of any portion of space-time, however small, large, irregular, or discontinuous. I have been wont to view events simply as physical objects in this sense. If Sebastian chews gum all the way across Bologna, and no longer, that event of his chewing and that event of his walking have been for me identical; they take up the same place-time. [Quine 1985, p. 167]

[212] This objection applies not only to perdurantism as a view of persistence, but to B-theory as well: if all the truths about the poker are tenseless truths describing how it is at different times and different places, nothing in what is true of the poker changes over time—hence there is no change on such a view. The objection goes back, in this version, to McTaggart [1927, ch. XXXIII, §§315–16], who was objecting to Russell's [1903, §442] B-theoretic understanding of change. See also Geach [1972, §10.2], Mellor [1981, pp. 110–11], and Simons [1987, pp. 134–7].

One cannot but get the feeling that, in the first-gear metaphysician's block universe, nothing happens. Everything is just sitting there, occupying its four-dimensional region within the big container that is reality—much like typical first-gear entities, such as numbers and triangles, are just sitting there, occupying their proper places amongst other numbers or geometrical entities.

In short, then, first-gear metaphysics deals with all the time-related notions, such as persistence, change, tense, and the like, by understanding them in terms of just one more dimension over which objects extend—the temporal dimension. We should note, however, that even though on this Lewisian first-gear view on time, perdurantism, B-theory and eternalism form a kind of package deal together with classical extensional mereology, strictly speaking the ingredients of this package deal are independent (to differing degrees). It has to be admitted, however, that taking, e.g., the combination eternalism *cum* endurantism or A-theory *cum* perdurantism results in rather odd views, which is due, we would suggest, to the fact that the *motivation* for all of the mentioned ingredients stems largely from a preference for first-gear thinking (as we are illustrating in this section), while their opposites are typically second-gear views (but see §6.2 for qualification of this statement).[213]

5.1.2 Modality

We observed, in §4.2, that typical first-gear entities, such as numbers, sets, and triangles, do not admit of *contingency*—it is no contingent matter that a given triangle is right-angled, if it is. But the way things are, in the big spatiotemporal manifold, presumably *is* contingent (at least to some extent). Luckily, for the first-gear metaphysician, contingency can be simulated rather easily, as we also observed in §4.2, by devising a suitable range of mathematical entities—models for a predicate-logical language, say—that differ in which predicate applies to which entity (or entities) of the domain. Now, given the idea of reality as being a spatially and temporally comprehensive four-dimensional manifold, we can devise a suitable first-gear simulation of contingency as follows. Consider a range of different distributions of basic properties over the constituent points of such a four-dimensional manifold. We designate one of those possible property distributions, the one supposedly corresponding to the actual distribution of basic properties over all of reality, as the 'actual world'. The others are 'possible worlds'. And now we may indeed say that the

[213] The combination of eternalism and endurantism has been defended by, e.g., Mellor [1981, 1998] and Johnston [1987]. The combination of presentism and perdurantism has been considered by, e.g., Benovsky [2009b], who rejects it, however.

'actual' distribution of properties over the four-dimensional manifold is a contingent matter—different distributions are 'possible', after all. There is no need, moreover, to move beyond the forms of predication typical of the first conceptual gear: simple feature predication is all we need to ascribe the relevant properties to points in the manifold. We will see shortly that the way modality is incorporated within Lewis's first-gear metaphysical program comes down to just such a simulation.

The Lewisian version of a first-gear metaphysical view we have developed so far purports to provide all one needs to account for everything whatsoever: we have a spatiotemporal manifold which instantiates, at each of its points or regions, certain properties, resulting in an entirely contingent arrangement of qualities. Within this manifold, we find objects composed according to the principles of classical mereology, and that's all; these provide the materials for understanding every phenomenon in reality. Lewis has dubbed this comprehensive metaphysical view *Humean supervenience*; he provides an almost plastic-pictorial characterization of it:

> Humean supervenience ... is the doctrine that all there is to the world is a vast mosaic of local matters of particular fact, just one little thing and then another. ... We have geometry: a system of external relations of spatiotemporal distance between points. Maybe points of space-time itself, maybe point-sized bits of matter or aether or fields, maybe both. And at those points we have local qualities: perfectly natural intrinsic properties which need nothing bigger than a point at which to be instantiated.[214] For short: we have an arrangement of qualities. And that is all. There is no difference without difference in the arrangement of qualities. All else supervenes on that. [Lewis 1986b, pp. ix–x]

Given this understanding of reality, and the means we just sketched for viewing it as being composed of matters of fact that are *contingent*, the path towards incorporating modality into this picture starts to become visible. Up until now, we did not find any reason to consider modality at all—the picture on which the thesis of Humean supervenience rests is an amodal picture. In fact, it is telling that for our sketch of first-gear metaphysical views so far we have not found any need to employ modal notions at all: that is fully in line with the peculiarities of the first conceptual gear. Although we should repeat, as argued in §2.3 and in §4.2, that there nevertheless is essentiality, which grounds its own (individuative) modality: the conception of physical objects involved is based on a mereological or spatiotemporal principle of individuation, which generates essential possession of spatiotemporal boundaries as well as essential part-whole relations for every individual such object.

[214] We should not misunderstand Lewis here: he is not defending the physics that seems to be presupposed by his present formulation. In a later article he makes explicit that he defends 'the *philosophical* tenability of Humean Supervenience,' and that his defense 'can doubtless be adapted to whatever better supervenience thesis may emerge from better physics' [Lewis 1994, p. 226].

Apart from the very thought that the particular matters of fact composing reality should be 'contingent', and apart from the fact that the notion of supervenience is itself a modal notion, it is clear that our daily dealings with the world are full of modality too: cats can catch mice, Fido can fetch the newspaper, tables cannot be placed on the surface of the sun, etc. Hence the first-gear metaphysician needs to come up with a story as to its meaning—just as in the case of temporality and change. Lewis has answered this challenge by espousing his notorious doctrine of 'modal realism'—which has a most unfitting name, since it is intended to be precisely a *reduction* of the modal to the non-modal.[215] According to this view, the modal features of some thing do not depend merely on that thing itself—that would be intrinsic, 'real' modality—but rather depends on how that thing relates to its counterparts in other possible worlds. The strategy is one of reducing modality to generality: something's possibly being F is reduced to *there being* a counterpart which is F, while something's necessarily being F is reduced to *all* its counterparts being F. The various possible worlds can be thought of as the myriad ways in which 'perfectly natural intrinsic properties' can be distributed over all of space-time, yielding alternatives to the Humean mosaic that is our actual world—much like the different distributions of properties over the four-dimensional manifold we considered above (see p. 164).

Now, famously, Kripke introduced the following example in his critical discussion of modal realism: Humphrey's possible victory in the 1968 elections consists simply in his having a counterpart who indeed was victorious in (the counterpart of) that unfortunate 1968 election. Kripke objects that Humphrey 'could not care less whether someone else, no matter how much resembling him, would have been victorious in another possible world' [Kripke 1972, p. 45, fn. 13], whereas he does take interest in the fact that *he himself* could have won the elections, if only he had done thus-and-so. Two objections can be extracted from this Kripkean quote: first, that his alleged counterparts are irrelevant to Humphrey's having certain modal features, and second, that Humphrey doesn't *care* about his counterparts but does care about himself. Again, let us listen to Sider in defense of the Lewisian view:

> [A]ccording to counterpart theory, the property of possibly winning *is* the property of having a counterpart who wins. Humphrey has a counterpart who wins, and so Humphrey himself (pound, stamp!) might have won. . . .
> Humphrey takes different attitudes toward the properties *possibly winning* and *having a counterpart that wins*. So those are distinct properties. . . . [Reply:] this is

[215] The *locus classicus* for this view is [Lewis 1986a], but the groundwork for the position was already laid in Lewis [1968, 1973b].

> just the paradox of analysis. A reasonable person can care about a property under one description ("possibly winning") while not caring about the same property under another description ("having a counterpart who wins"), provided it is not obvious that the descriptions pick out the same property. Correct analyses need not be obvious to competent language users. Obviousness may count for something, but theoretical virtues are important as well in determining which analyses we ought to accept. [Sider 2006b, p. 1–2]

As in the case of change, these objections, and the ensuing replies by Sider and other defenders of the Lewisian approach, nicely illustrate a clash between first-gear thinkers and higher-gear thinkers: those insisting on the first gear simply observe that the best way to make sense of notions that find their proper home within higher gears is by way of their preferred reductive analyses; while those in favor of the second conceptual gear will never be content to accept such analyses simply because the whole *point* of having, e.g., modality or temporality play a role in their metaphysical thought lies in adopting this higher gear, according to which time and modality are essentially different from their first-gear surrogates. The futility of engaging in a debate on what, e.g., modality is, without paying due attention to this radical difference in conceptual starting point, is nicely expressed by Sider's half-joking use of the phrase 'pound, stamp!'.

First-gear metaphysicians are likely to think of modality along the sketched lines, even if they do not go as far as Lewis in defending modal realism.[216] For our purposes, it is interesting to note that such approaches to modality fail to take into account the diversity of modal notions that we are able to distinguish if we give up on restricting ourselves to the first conceptual gear. Apart from the generic modality that is involved in conceptual truths, which we have called conceptual modality, we also found further modal notions in connection with our introduction of second-gear temporality and causality in the previous chapter (§4.3): we distinguished between tenseless, in-principle power ascriptions and tensed, situated power ascriptions, where the former corresponds to generic causal possibilities while the latter corresponds to local, real possibilities. The richness of this second-gear perspective on modality is totally absent from first-gear treatments of modality such as Lewis's. Instead, we find a series of ingenuous and seemingly rich, but actually austere simulations.

All in all, then, the metaphysical picture that Lewis and his followers construct by applying the first conceptual gear to all of reality consists of elements that are

[216] Quine, for instance, famously mounted a strong campaign *against* modality—see, e.g., Quine [1943, 1947, 1948, 1953b,c, 1976b, 1991]. On the other hand, Williamson [1999, 2010, 2013], who defends 'necessitism', the view that necessarily, everything necessarily exists, can easily be seen to fall into the category of first-gear metaphysicians without either reducing or rejecting modality. First-gear metaphysics comes in many guises.

virtually made for each other, that complement one another in quite natural and coherent ways to yield first-gear simulations of all notions one would at first sight locate within the second conceptual gear.[217] The appeal of such a picture lies precisely in its mathematical nature—witness Sider's following, very telling remark ('four-dimensionalism' encompasses eternalism, B-theory and perdurantism):

> It is easy to feel ... an intellectual joy in contemplating a theory so elegant and beautiful as four-dimensionalism, and it is tempting to accept the theory simply on this basis .... [Sider 2001, p. 74]

And it is interesting to compare this remark with a remark from David Wiggins, who defends second-gear notions against such first-gear simulations:

> An idea like that[218] could not even occur to one with the good fortune to be innocent of classical extensional mereology. [Wiggins 2001, p. 166]

This concludes our survey of first-gear metaphysics. In the next section, we deepen this survey by engaging in an in-depth discussion of one topic that not only further illustrates first-gear metaphysics but also makes clear its connection with Humeanism: laws of nature.

5.2 Necessary Connections and Laws of Nature

We have said that first-gear metaphysics is Humeanism, but, apart from the doctrine which Lewis calls 'Humean supervenience', we have not seen much Humeanism yet. This is because we have so far ignored the topics that bring to mind typically Humean views, viz., causality and laws of nature—two basic concepts of the second conceptual gear (as developed in §4.3). As we will at the end of the present section, it turns out that, on reflection, Humeanism and first-gear metaphysics do not exactly coincide—rather, Humeanism allows for *at most* first-gear metaphysics.

But we should not get ahead of things. Humeanism with regard to causality and laws of nature starts with strong suspicions about what Hume called 'necessary connections', and ends with analyses of causality and of laws of nature that do not incorporate such necessary connections. The general idea is familiar enough: in reality, it is thought, things just happen, there is no 'causal glue' that ties different events together, no laws that 'govern' the behavior of objects. This is the idea

[217] We will see how it extends to laws of nature and causality in §5.2 and §6.1, respectively.

[218] Wiggins is here talking about the idea that there are as many mountains having the same peak as there are rival determinations of boundaries around it—but this context is irrelevant here; the quote is merely meant to illustrate second-gear dislike of first-gear metaphysics.

that Lewis then reworked into his paradigmatically first-gear thesis of Humean supervenience—see the quote we gave on p. 165 above. Below, we develop the idea of such a Humean mosaic of loose and separate matters of fact with some care, starting from Hume himself. We then go on to discuss Humean approaches to laws of nature, in order to show that such approaches indeed require a first-gear metaphysical understanding of the underlying mosaic, that is, an understanding much like the Lewisian one we have sketched in the previous section.

We have chosen to focus on laws of nature in this chapter, and to discuss causality in the next chapter (see §6.1). We find laws of nature, because of their generality, to be more suitable for purposes of illustrating the way in which second-gear concepts can be simulated by using first-gear thought, while we take causality to be more suitable for illustrating the second conceptual gear itself, as the shift in recent decades from a purely categorical understanding of causality as a relation between events (i.e., in first-gear terms) to a more dynamic, powers-based understanding of causality makes clear (recall our powers-involving description of second-gear thought in §4.3).[219]

5.2.1 On the Idea of a Humean Mosaic

Let us start by turning to Hume himself. Hume famously complained that '[t]here are no ideas, which occur in metaphysics, more obscure and uncertain, than those of *power, force, energy,* or *necessary connexion*' [Hume 2007/1748, VII.1, p. 45].[220] On Hume's empiricist view, all ideas (concepts) are created by the mind on the basis either of impressions from the external senses or on the basis of internal sentiments.[221] Hence he proposes to enlighten us about these purportedly obscure ideas by way of the following method:

> By what invention can we throw light upon these ideas, and render them altogether precise and determinate to our intellectual view? Produce the impressions or original sentiments, from which the ideas are copied. [Hume 2007/1748, VII.1, p. 46]

Hume goes on to put to use this 'new microscope or species of optics' [ibid, VII.1, p. 46], and concludes that the idea of necessary connection does not come to us through the external senses:

[219] Indeed, defenders of second-gear thought have gone so far as to reject the very idea of laws of nature, meaning to discard first-gear conceptions of these. Such is our reading of, for instance, Cartwright [1984].

[220] Interestingly, all of the ideas Hume lists here as being obscure and uncertain are paradigmatically second-gear ideas—as we have indicated in §4.3.

[221] This is, of course, the aspect of the Modern Picture we called *Mind-Dependence* in §1.2.

> When we look about us towards external objects, and consider the operation of causes, we are never able, in a single instance, to discover any power or necessary connexion; any quality, which binds the effect to the cause, and renders the one an infallible consequence of the other. We only find, that the one does actually, in fact, follow the other. [Hume 2007/1748, VII.1, p. 46]

Hence Hume goes on to look for internal sentiments on which the ideas of cause and necessary connection are grounded, and indeed claims to find such sentiments—resulting in his view that the idea of causation is derived from the mind's habit to expect the effect whenever the cause is observed (roughly). But we do not need to delve any deeper into Hume's own conception of causality here.

Hume's view of what the external senses deliver is the first version of what we call a *Humean mosaic*: it is just particular matters of fact that are entirely loose and separate. It consists of things being thus over here right now, things being so over there a little bit later, etc. All such particular matters of fact are independent from each other, there is no 'glue' that binds how things are over here now to how things are over there then, except for the purely external relations of spatial and temporal distance. In effect, the idea of such a Humean mosaic only allows for what we have called feature predication in §4.2: all local, particular matters of fact consist just in something's (or some things') having a certain property (or standing in a certain relation), where the properties and relations in question are purely categorical (i.e., non-modal; dispositions and the like—'power, force, energy'—are disallowed).

Now, in Hume's case, at least on our reading, it seems that this mosaic is to be understood in an epistemological way: the senses provide us with a stream of impressions none of which contains anything which somehow constrains other impressions. This, of course, later became the guiding thought of phenomenalism and of logical empiricism.[222] But this epistemological understanding is not the only way of endorsing a Humean mosaic: as we have seen already, Lewis defends a metaphysical version of it (see §5.1.1). On Lewis's view, the independent elements of which the mosaic is composed are not construed as impressions or as sense-data but rather as local, particular matters of fact within a spatiotemporal manifold. According to Lewis's doctrine of Humean supervenience, everything there is to say about reality can in the end be reduced to just such a metaphysical Humean mosaic.

Whether understood epistemologically or metaphysically, the idea of a Humean mosaic has played an important role in philosophy ever since Hume so vividly

[222] Hume himself can be taken to have started the phenomenalist tradition—see esp. Hume [2010/1740, Part IV, §§2–4], but probably Mill [1865, chs. XI–XIII] contains the first thorough statement of phenomenalism.

presented it to mankind. However, and especially in the empiricist tradition within analytic philosophy, Hume's search for 'internal sentiments' was discarded, and instead logical *reductions* of causal and nomic notions to patterns in the Humean mosaic were proposed. Contemporary philosophers of science often proudly remark that they are 'Humeans', meaning that they place themselves within this venerable tradition. It is to this tradition that we now turn in order to find out how such Humeanism relates to first-gear metaphysics. As said, we focus on laws of nature.

5.2.2 Humean approaches to Laws of Nature

The classical Humean approach to laws of nature takes them to be just regularities within the mosaic: it is a law that *A*s are *B* just if *A*s are always *B*. Given that laws of nature are often thought to incorporate some kind of necessity, viz., nomic or natural necessity, we here have another instance of the familiar reductive strategy of replacing modality by generality, a typical symptom of which is a reversal in the order of explanation: whereas intuitively, the laws are to explain the behavior of things, the Humean has to say that the behavior of things, taken together, explains the laws.[223] And, precisely because it ignores this modal aspect, the Humean take on laws of nature faces a typical problem: it lumps together those generalizations that are genuine laws with other, merely accidental generalizations. As we will see shortly, none of the many versions of Humeanism, neither the crude one just delineated nor the many sophisticated ones that result from attempts to improve the Humean approach, can be subscribed to without thereby also endorsing a Lewis-like conception of the underlying Humean mosaic—that is, without endorsing first-gear metaphysics.

Early in the twentieth century, philosophers of science proposed various strategies for distinguishing the law-like generalizations from the accidental ones. Usually, these strategies involved pointing to peculiarities in the way we come to believe generalizations, or to the role such generalizations play in our theories, or to restrictions on the content of the generalizations, such as not mentioning any specific places or times. For instance, Braithwaite [1927, 1928] proposed to understand the laws to be those generalizations that are 'believed for reasons that are not logically demonstrative' [Braithwaite 1927, p. 467], that is, which are not deductively derived from particular matters of fact.[224]

[223] Recall the reversal in the order of explanation we observed in our discussion of the reductive, Tarskian approach to logical validity in §3.2, p. 92.

[224] A more recent defender of such an orthodox Humean view on laws of nature, Swartz [1985], summa-

Ramsey was dissatisfied with this proposal of Braithwaite's, as well as with other such proposals: it seemed to him that a generalization's being a law should be independent from how we happen to achieve knowledge of it or currently think of it. Instead, he sketched an alternative:

> What is it then that is true of universals [of law] and not of [universals of fact]? We have seen that it is not their spatio-temporal indifference, not that they are believed .... Also the difference would persist if we knew everything.
> This last point gives us the clue; even if we knew everything, we should still want to systematize our knowledge as a deductive system, and the general axioms in that system would be the fundamental laws of nature. [Ramsey 1928, p. 131]

Almost half a century later, Lewis developed this idea into what is now widely perceived to be the most promising candidate of a Humean account of laws of nature—the so-called Best Systems or Mill-Ramsey-Lewis account[225]:

> We can restate Ramsey's 1928 theory of lawhood as follows: a contingent generalization is a law of nature if and only if it appears as a theorem (or axiom) in each of the true deductive systems that achieves a best combination of simplicity and strength. [Lewis 1973b, p. 73]

It is clear that the idea of a Humean mosaic plays a crucial role in this theory of lawhood. Reality consists of just a large mosaic of particular matters of fact which displays certain patterns that enable concise summaries of the mosaic. Such summaries can be phrased, so the thought goes, as deductive systems with certain axioms—the candidate Best Systems—which can then be compared as to how well they score on the balance between simplicity and strength.[226] Simplicity is cashed out, for example, in terms of some complexity metric, whereas strength is understood in terms of the amount of information embodied in the axioms—stronger systems say more about the entire mosaic.[227]

As Lewis observes, it is not even required, on the Best Systems approach, to restrict the laws to general truths: if it improves the overall balance of strength and

---

rizes his position as follows: 'there are a number (an infinite number, perhaps) of true universal, material conditional statements (propositions), all of whose terms are perfectly descriptive, that is, make no reference to any particular time, place, person, or thing in the world. The Regularist is content to allow that all these true universal, material conditionals are physical laws ....' [p. 101].

[225] Ramsey himself abandoned his own proposal already one year later because 'it is impossible to know everything and organize it in a deductive system' [Ramsey 1929, p. 138].

[226] Apart from Lewis himself, defenders of the Best Systems approach, in one or another version, include Albert [2001], Beebee [2000], Dunn [2011], Earman and Roberts [2005a,b], Loewer [2007b, 2009, 2012], J. Cohen and Callender [2009], and Callender and J. Cohen [2010].

[227] One might worry about the status of these norms of simplicity and strength: do they not incorporate a psychological aspect into what it is to be a law? Lewis thinks not, as long as 'nature is kind enough' to convincingly make one system the absolute best one—see Lewis [1994, p. 479].

simplicity, a truth about particular aspects of the entire mosaic will be regarded as a law as well.[228] What Lewis does require, however, is that we run the Best Systems competition on the basis of a predicate-logical language containing a certain privileged set of predicates. Without that restriction, so he rightly thinks, the whole project becomes trivial: we could simply choose a language, say, with a primitive predicate P which is stipulated to apply to an object just if it is part of a reality which is composed of exactly those matters of fact that make up the actual mosaic. The generalization 'Everything is P' would then easily win the competition because it is maximally strong and (arguably) maximally simple—but also very uninteresting. Now, Lewis's way of restricting the language for which we run the Best Systems machinery is by endorsing a privileged set of 'perfectly natural properties': those are the fundamental properties that are instantiated by the basic building blocks of the universe—e.g., by the space-time points.[229]

In fact, the privileged set of natural properties is already required to arrive at a Humean mosaic in the first place: an abundant construal of properties, on which *any* set of objects amounts to a property, trivializes the whole idea of (regularities within) a mosaic. In terms of our colorful analogy: if the set of yellow dots within the mosaic is no more special than any other random set of dots within the mosaic, there is no way of recognizing patters in it, except by first picking out a few of the many sets of dots to be treated as special. And in that case, it will be more appropriate to speak of *drawing* patterns into the mosaic than of *recognizing* patterns in it.[230]

Hence, the (predicate-logical) language on which the Best Systems competition is to be based should include primitive predicate symbols for just the perfectly natural properties; contrived ones like our P are not allowed. In this way, the Best Systems theory also avoids the allegedly impossible task of having to compare the balance of simplicity and strength for systems couched in different vocabularies.[231]

Following Lewis, many have found it natural to identify this set of perfectly natural properties with the fundamental properties of physics.[232] Yet this creates a

[228] Lewis does not think that 'there are laws that essentially mention Smith's garden, the center of the earth or of the universe, or even the Big Bang', but still, 'such laws ought not to be excluded *a priori*' [Lewis 1986b, p. 123]. Albert [2001] and Loewer [1996, 2007a] make use of this option in arguing that we should include the 'Past Hypothesis' (viz., that the universe started out with a very low entropy state) in the Best System. See Frisch [2011] for critical discussion.

[229] See Lewis [1983, pp. 366–8] and Lewis [1986b, pp. 123–4]. Lewis envisages three options for privileging natural properties: by means of a primitive predicate of naturalness; by providing a reductive analysis in terms of universals; or in terms of tropes. But he does not take a stand: 'the honors are roughly even, and remain undecided' [Lewis 1986a, p. 64].

[230] But see Taylor [1993] for such a theory-relative (and hence anti-realist) construal of naturalness.

[231] This point is developed in detail by J. Cohen and Callender [2009, pp. 5–8].

[232] For a clear statement of just how Lewis intends this identification to be read, see Lewis [1994,

problem with regard to the laws of the other sciences, unless one is prepared to accept that only the laws of physics are 'true' laws of nature. But not everyone is prepared to go that way; hence an amended version of the Best Systems approach has recently been put forth: the so-called 'Better Best Systems' approach.[233] On this approach, each science has its own laws resulting from a Best Systems competition based on the specific vocabulary of the science in question. It is thought that this amendment to the Best Systems approach makes room for autonomous special sciences.[234]

5.2.3 The Underlying Mosaic

We now have a reasonably diverse menu of Humean options on the table. What ties them together is, indeed, their 'Humeanism': the laws do not govern what happens in the world, but are mere summaries of what happens—that is, summaries of the entire Humean mosaic. Each of these options builds on the idea of a Humean mosaic that we inherited from Hume. Now, what does the mosaic have to look like in order to provide a useful foundation for this type of approach? Most of the defenders of a Humean theory of laws do not develop any detailed conception of this underlying mosaic—except, of course, for Lewis, as we have seen. In fact, Callender and Cohen regard it to be a virtue of their Better Best Systems view that it is *not* committed to Lewisian metaphysics. Instead, they endorse 'explosive realism':

> [T]he world permits possibly infinitely many distinct carvings up into kinds, each equally good from the perspective of nature itself, but differentially congenial and significant to us given the kinds of creatures we are, perceptual apparatus we have, and (potentially variable) matters we care about. [J. Cohen and Callender 2009, p. 22][235]

It is not at all clear, however, what such an explosive realism implies for the conception of the underlying mosaic. The thought seems to be that it does not matter for 'the world', that is, for the Humean mosaic, whether or not we approach it with just a limited stock of perfectly natural properties in mind, or rather with more promiscuous

esp. p. 474].

[233] See Schrenk [2006] and J. Cohen and Callender [2009] and Callender and J. Cohen [2010]. The label 'Better Best Systems' is Callender and Cohen's.

[234] Interesting questions can be asked about the proposal. For example, in biology most of the vocabulary from physics is also used, but it seems possible that some law of physics, resulting from a physical vocabulary based Best Systems competition, will not be a law in biology.

[235] Recall our brief reflection on this type of metaphysical view at the end of §1.1.2. Cohen and Callender borrow explosive realism from Sosa [1993, 1999], Taylor [1993], and Dorr [2005], and Eklund [2008], amongst others.

'carving tools'.[236] Lewis, on this thought, is unduly restrictive in demanding his favored set of perfectly natural properties to be somehow metaphysically privileged. However, by focusing solely on the *properties* that are or are not 'allowed', this way of thinking leaves out one important issue: the *objects* that occupy the mosaic.

Recall that, on Lewis's view, the mosaic is populated with space-time points (or occupants of space-time points), and with sums of these. The only 'natural kind' that plays a role on this ground-level of the mosaic is the kind 'physical object' which is defined as just the sum of its point-sized parts (or, alternatively, as just the contents of the spatiotemporal region it occupies). We argued in §2.3 that, by application of the essentialist inference, Lewis is therefore committed to mereological and/or spatiotemporal essentialism. In the previous section, we have seen that Lewis reintroduces the modal flexibility we associate with the objects that occupy reality by way of counterpart theory—but the mosaic *itself* is not touched by this fact: together with all the other possible worlds, it forms the reductive base for his counterpart-theoretic analysis of modality.

The proponents of explosive realism equivocate on how exactly their 'explosion' is to be understood: does the mosaic *itself* contain all the various kinds of objects to which the possibly infinite ways of carving endorsed by the explosive realist give rise, or is it rather that the mosaic 'permits' (as Cohen and Callender are saying, see the above quote) all those different carvings into objects?

Consider the first option. The mosaic, on this proposal, contains not just physical objects in Lewis's sense, but also things like rabbits, tables, protein molecules, mountains, oak trees, and many, many unheard-of things. Of course, Lewis's mosaic can be said to contain all of these as well. But it is important to note that they do not play a role in Lewis's *conception* of the mosaic: there are rabbits in the mosaic because there are suitable physical objects (in his sense) that display a leporine distribution of qualities and stand in appropriate relations to their counterparts in certain other possible worlds. That is to say, for Lewis rabbits are not ontologically fundamental, in the sense we developed in §2.2.1: to be a rabbit just is to be a physical object satisfying the mentioned conditions. Hence, strictly speaking, the Lewisian mosaic is not *composed* of things like rabbits, but rather *grounds* them by containing physical objects, to which the rabbits reduce. Or, as Lewis would say, the rabbits supervene on the distribution of qualities.

[236] Although we are not particularly fond of the carving metaphor, we will indulge in it so long as we are discussing philosophers who are wont to using it.

Now, what would it mean to say that the mosaic *itself* is composed of rabbits and the like, as we are considering on behalf of the explosive realist? Put in terms of §2.2.1, this would mean giving up on a reductive approach to *all* of these kinds. For, as soon as we reductively explain the presence of rabbits in terms of some appropriate things' fulfilling certain conditions, as on the Lewisian view, we are thereby removing them from our conception of the mosaic. The result is thus an incredibly wealthy catalogue of ontologically fundamental things. That is not all by itself a problem, of course, but it does create a tension with the very idea of a Humean mosaic, which is supposed to be free of necessary connections. Let us see why.

Rabbits don't simply pop into existence, they come from their parents. A rabbit only stays alive as long as it takes in suitable nourishment, meaning that in its past there has to have been enough such nourishment in its proximity. Rabbits don't turn into butterflies, meaning that in the future it will not suddenly transform into one. It is clear, then, that if something here-now is a rabbit, this has lots of implications as to how things are at other places and times: there have to have been parents, nourishment, absence of conditions too extreme for rabbits to survive, etc. In other words, if something here-now in the mosaic is a rabbit, and rabbits are ontologically fundamental things, then the mosaic contains 'necessary connections'. Nor does it make a difference to change the example to one drawn from physics: electrons, like rabbits, don't pop into existence either; they don't suddenly change from being negatively charged to being positively charged; when shot into cloud chambers they do not trace out the word 'electron', etc. These necessary connections make the Humean approach to laws of nature entirely superfluous: we can simply list the necessary connections that we already have within the mosaic.[237]

The problem lies with the conceptual truths associated with the concept *rabbit* (or *electron*), so, one might think, the remedy is easy: we just strip our conception of a rabbit (or electron) from all of these conceptual truths, which leaves us with a purely qualitative residue of rabbithood (or electronhood—though one might doubt whether anything at all is left over in this case). And then we can say that the mosaic contains rabbits (or electrons) in this watered-down sense. On reflection, however, this proposal comes down to a reduction of rabbits and electrons to something which is ontologically more fundamental, much like Lewis's proposed reduction to physical objects in his sense—in other words, we have nothing new.

[237] We may of course go on to systematize these connections in a way much like the Best Systems approach proposes, but it is clear that this does not even come close to being a Humean approach to laws of nature.

Consider, now, the second proposal of construing a mosaic based on explosive realism: the mosaic, on this second construal, does not *contain* all the different objects, as in our first proposal, but rather *permits* being carved up into all those different kinds of objects (as Callender and Cohen put it). One very straightforward way of giving meat to this idea is Lewis's: his mosaic contains just physical objects (in his sense), and these indeed permit to be 'carved up' in myriad different ways. One may define a kind of object *K* by just listing random qualitative and modal criteria; and if there are any physical objects (in Lewis's sense) fulfilling these criteria, then there are *K*'s. This works for rabbits (as Lewis thinks of them) just as well as it works for unheard-of kinds of things. But now explosive realism seems, again, to be nothing new: we are back where we started, with Lewis's metaphysics.

A genuinely new option is in the offing once we decide to treat Lewis's physical objects, which form the reductive base which permits all the kinds of carvings, as *itself* just the result of a carving operation. That is to say, we stop assuming a certain bottom-level realm of things to which everything can be reduced. By doing so, we suddenly find ourselves having left the camp of metaphysical realism altogether: we stop asking for an account of what is ontologically fundamental, we simply say that every carving is the result of our applying concepts to something which also permits being carved up differently. No carving is ontologically fundamental anymore, as it is on Lewis's account—or rather, ontological fundamentality becomes a matter of stipulation.

It might be rather surprising that explosive realism leads us straight to *anti*-realism. But this is not really so: we have just discovered that what is *called* 'explosive realism' is, at least on one reading of it, realism only by name, just as Putnam's 'internal realism' and Blackburn's 'quasi-realism' are realisms only by name (see §1.1.2).[238] What is more worrying, especially for proponents of a Humean theory of laws, is that the entire project now starts to break down. That is because the mosaic as it is independently of our carving it up has now become intractable—in contrast to Lewis's view, on which there is a way of considering the mosaic independently from our way of carving it up, viz., in terms of physical objects (in his sense). The mosaic, independently from our carving, is just an 'amorphous lump', to repeat Dummett's phrase (see, again, §1.1.2), and the only way of saying something about it—for instance, that it exhibits certain regularities—is by first applying our carving tools.

[238] Aristotle would have said that such realisms are merely homonymously called 'realism'. See the opening sentence of the *Categories*: 'When things have only a name in common and the definition of being which corresponds to the name is different, they are called *homonymous*.' [Aristotle 1963, p. 3, 1a1–2].

In order to see why this makes the Humean approach to laws of nature break down, we need to return to Lewis's view on rabbits. We saw that he assumes that rabbits reduce, ontologically speaking, to his physical objects. A physical object is a rabbit just if it displays a suitable arrangement of qualities *and* also has the right modal properties, which are spelled out in terms of its counterparts. Now, the counterpart relation itself depends on laws of nature, on Lewis's view: the worlds in which our laws of nature are true are closer than the other worlds. They form the 'inner sphere'. Hence the basis on which the laws of nature are determined cannot involve rabbits, or any other objects that involve modality. That is to say, the mosaic depends solely on the distribution of qualities over physical objects, in Lewis's sense. His conception of physical objects, together with his array of purely categorical, 'perfectly natural' properties, ensures that there is a decent basis for the Best Systems machinery to operate on. The carving can only proceed once we already have the laws of nature. Lewis's metaphysical system is, indeed, like a carefully constructed multistage rocket.

By contrast, our anti-realist reading of explosive realism leaves no room for such bootstrapping procedures. Given that no way of carving is ontologically fundamental, it is simply arbitrary to *first* carve into spatiotemporally individuated physical objects, *then* run a Best Systems machinery, and *then* do the rest of the carving. We could just as well simply carve directly into rabbits and the like, and then have the conceptual connections we thus introduce play the role of laws of nature. The resulting view would be congenial to the projectivist theory of laws of nature that Nelson Goodman and others promoted.[239] Of course, the point of the Humean, regularity-based account was that the regularities are objective, mind-independent, and the resulting account of laws of nature therefore 'admirably realist when compared against projectivism' [J. Cohen and Callender 2009, p. 2].

We can now see that in order to arrive at such an objective and mind-independent construal of the mosaic, one needs to endorse some kind of positive conception of what that mosaic is fundamentally like. We have used Lewis's construal of the mosaic as an example, but strictly speaking this is not the only option. What matters is that the physical objects that occupy the mosaic (as well as the properties that are instantiated by those objects) are construed in a modally flat way—e.g., mereologically, spatiotemporally, or in some other way. And that is just what characterizes first-gear metaphysics. In conclusion, then, we may say that the Humean approach to laws of nature only makes sense against the background of first-gear metaphysics.

[239] See, e.g., Goodman [1955], Ayer [1956b], and Ward [2002].

5.2.4 Humeanism and First-Gear Metaphysics

The misgivings about 'necessary connections' that are so typical of Humeanism are, in the end, symptoms of a very deeply rooted suspicion towards *conceptual* connections in general: conceptual connections are mere connections amongst ideas; what could they possibly tell us about reality? To say that they capture real 'necessary connections' out there is just to project what belongs within the mind to the outside world. That is why the Humean is drawn towards metaphysical views on which these are as thin as we can possibly get—that is, towards first-gear thinking.

Yet we have here also reached the limits of our assimilation of Humeanism to first-gear thinking. It is more appropriate to say that, *if at all*, the Humean ends up endorsing first-gear metaphysics—the other option being anti-realism. One could of course urge that true Humeanism entails rejection of metaphysics *tout court*, in which case only the anti-realist type of Humeanism would count as truly Humean. Such terminological issues are not our concern here, but they do point towards the following potentially illuminating considerations.

If it is reification of conceptual connections that is bothering the Humean, there is considerable pressure for him to adopt anti-realism. Hume says that, if we look closely, we never *see* any necessary connections. By extension, we can argue that we never *see* any succession in time or any contiguity in space; we only see the sensory qualities, the spatial and temporal relations between them are not *themselves* such qualities. Indeed, even the *difference* and the *resemblance* between different sensory impressions are not themselves impressions. The Humean is under pressure, then, to say that all of these are the result of what the mind does with the given impressions, much like the mind, on Hume's view, creates an expectation of the effect whenever the cause is present. In the end, then, this type of suspicion towards the conceptual naturally tends to end in a broadly Kantian view on which *no* conceptual determinations of reality-itself can be given.[240]

The first-gear metaphysician, on the other hand, is a metaphysical realist: he takes our first-gear ways of thinking to capture the true structure of reality-itself, while the rest of our conceptual repertoire is to be explained in terms of such first-gear concepts. That is, his suspicion towards conceptual connections is restricted to just those conceptual connections that go beyond the first gear. As we have seen, Lewis, and especially Sider, are typical first-gear thinkers in just this sense.[241]

[240] We should repeat that we do not mean to endorse any specific interpretation of Kant himself with this statement—-see fn. 46 on p. 31.

[241] Sider illustrates our claim especially in his last book, called *Writing the Book of the World* [2012]. One

5.3 Concluding Remarks

We have seen how the application of first-gear forms of thought to reality as a whole naturally leads to a quite specific range of mutually supporting metaphysical views, most of which have been brilliantly defended by David Lewis: perdurantism, classical extensional mereology, Humean supervenience, modal reductionism. We then moved on to argue that there is a close relation between these metaphysical views and Humeanism in general. To that effect, we focused on Humean approaches to laws of nature.

Thus, our purpose in discussing Lewisian metaphysics and Humeanism about laws of nature has not been to critically assess them in the usual sense, but rather to illustrate the kind of metaphysical view one arrives at when sticking to the first conceptual gear across the metaphysical board. Hence our remarks, at several important junctures, on symptomatic clashes between first- and second-gear thinking as they arise within the metaphysical debates we have touched upon.

In the next chapter we will focus on such clashes, especially as they arise over questions regarding causality (§6.1) and time (§6.2). Where the present chapter has been, to some extent, descriptive in character, the next chapter will have a more discursive character: there we argue that moving towards second-gear thought only makes sense if the move is made across the board, that is, if we leave Humeanism behind for good.

* * *

could list Quine here as well, but his views are more difficult to assess. For what it is worth, we take Quine to be a metaphysical anti-realist in our sense, who nevertheless demands that we restrict ourselves to first-gear thinking insofar as we wish to be 'scientifically respectable'.

Chapter 6

The Second Gear: Anti-Humeanism

In this chapter, we show that dissatisfaction with Humean approaches to issues in metaphysics and the philosophy of science is to be understood as a demand for shifting from the first to the second conceptual gear. In particular, we look at dissatisfaction with the way in which first-gear thinking deals with two of the most central second-gear concepts: causation and time. Such dissatisfaction has yielded numerous views on causation and time that diverge from the more typical first-gear accounts. We argue that such divergence is of no use, as long as the basic way of thinking underlying the views in question is still limited to first-gear forms of thought, and we indicate what it takes, in both cases, to genuinely shift from the first to the second conceptual gear.

Thus we complete our investigation of the ways in which first- and second-gear thought shape metaphysical views, and of how exactly these views stand in fundamental opposition to each other.

We introduced the forms of thought that are typical of the three conceptual gears we discuss in the second part of this essay in chapter 4, and we argued in the previous chapter that first-gear metaphysics, which sticks to first-gear forms of thought across the metaphysical board, leads to Humeanism in general and to a metaphysical outlook much like Lewis's in particular.

By way of orientation, let us repeat the general scheme of our three conceptual gears here:

	First gear	**Second gear**	**Third gear**
form of predication:	tenseless	tensed/aspected	normative
applies to:	entity	object	organism
predicates involved:	feature	state/process	life-process
sortals involved:	type	natural kind	life-form
mode of being:	to exist	to persist	to live
engaged in:	—	process	life-process
characteristic:	geometry	causality/temporality	teleology/normativity
	Humeanism	*anti-Humeanism*	

A lot of what is going on in contemporary metaphysics as well as in certain parts of the philosophy of science can be understood as a clash between first-gear and second-gear thinking. We already saw that this is so in our description of first-gear metaphysics in the previous chapter: interesting things happen, philosophically speaking, just where the first-gear metaphysician tries to analyze second-gear concepts in first-gear terms, and we have learned a lot about first-gear thinking (and its relation to Humeanism) by looking closely at such attempts. Conversely, we can learn a lot about second-gear metaphysics by examining the ways in which second-gear metaphysicians try to resist such first-gear analyses. Given the predominance of first-gear (and Humean) thinking in the history of analytic philosophy, it is not surprising that such resistance towards first-gear analyses not always results in a truly second-gear metaphysical view: it is not easy to resist the tendency to underpin one's views with first-gear principles and assumptions. Yet insofar as one fails to resist this tendency, the resulting view collapses into a mere alternative first-gear analysis.

We illustrate this visceral entanglement in first-gear forms of thought by looking, in §6.1 below, at one very typical point of divergence between first- and second-gear thought: whether or not causality involves something like production, something metaphysically robust that goes beyond mere 'constant conjunction', to use Hume's famous phrase. The second-gear metaphysician urges that it does, while the first-gear metaphysician, following Hume, denies this, as in regularity theories of causation, or tries to reduce the relevant notion of 'production' to something unproblematic from a first-gear point of view, as in counterfactual analyses of causation. Our discussions on this point continue the work on laws of nature we have already done in the previous chapter (§5.2), although here we are no longer interested in what the Humean is committed to, metaphysically, but rather in what the anti-Humean is committed to. Our prime example in this section is David Armstrong, whose metaphysical view constitutes an interesting case of second-gear ideas incorporated

into a basically first-gear setting. We argue that what the anti-Humean is aiming at can only be successfully achieved once he is willing and able to get rid of the first-gear habits of thought that have so strongly (but by no means exclusively) shaped analytic philosophy. If so, he will end up endorsing something like the picture of second-gear thought we have sketched in §4.3.

To our discussion of causality, which has a somewhat descriptive flavor, we add, in §6.2 below, a more discursive investigation into the nature of time. Time constitutes another very typical point of divergence between first- and second-gear thinking. The first-gear metaphysician typically conceives of time in eternalist, B-theoretic, and perdurantist terms, as we saw (see §5.1.1). Time, however, is a central concept of second-gear thought: within the second conceptual gear, time is dynamic in a way that cannot be captured in first-gear terms. It consists in the applicability of the second-gear form of predication, which is tensed and aspected. The second-gear concept of time is caught in a conceptual circle together with change, persistence, causation, tense, process, and the like, and every possible way of breaking into that circle results, at best, in a mere first-gear simulation of time as a static, eternalist time line. This is true even for certain versions of presentism, A-theory, and endurantism, even though these latter three views do form a crucial part of the second-gear concept of time.

Most anti-Humean metaphysicians fail to arrive at a truly second-gear metaphysical view, which is due largely to the habit of mixing various first-gear reductive ingredients into one's attempt at a second-gear metaphysical view (for instance, insisting on necessary connections of some sort while also adopting an eternalist understanding of time, as Armstrong does). This results in lots of local debates within the confines of first-gear thinking. Such local debates are fruitless, for they fail to be the fundamental metaphysical clashes they purport to be: the real issue is whether or not the second conceptual gear is accepted. We conclude that a clearer understanding of how deeply the nexus of second-gear concepts differs from that of first-gear concepts allows for a quite definite rejection of all attempts to accept only part of the second conceptual gear supplemented with a first-gear analysis of the remainder.

6.1 Causation

We have extensively discussed the Humean approach to laws of nature in §5.2, where we also introduced the idea of a Humean mosaic, and argued that the only

sensible way for construing such a Humean mosaic is in first-gear terms (see §5.2.3). Traditional Humean accounts of causation rely on the same kind of approach that also underlie Humean accounts of laws of nature: the idea is, roughly, that *A* causes *B* just if *A* is followed by *B* and, in general, *A*s are *always* followed by *B*s—that is, just if the corresponding generalization is true (and, presumably, not just true, but also a law). Here, *A* and *B* are understood to refer to (kinds of) events, and events are to be understood as inhabiting the Humean mosaic. Causation is, thus, a relation between events in space-time, logically not unlike the relation 'bigger than' between physical objects.[242] Since the intuitive idea of causation involves some kind of production, which is understood to be a modal notion, we have again a fine example of reducing modality to generality, and, as we should expect, a corresponding change in the direction of explanation: whereas intuitively, causes are thought to explain the coming about of their effects, the Humean has to say that it is the 'constant conjunction' of cause and effect that explains their *being* cause and effect.[243]

Interestingly, and unlike in the case of laws of nature we discussed earlier, there is another thoroughly Humean approach to causation available. It is the analysis of causation in terms of counterfactuals made famous, again, by David Lewis. As is to be expected, this approach to causation does not move beyond first-gear metaphysics within the hands of Lewis. However, the idea has been picked up by lots of others who are not as strictly wedded to first-gear forms of thought as Lewis is—hence this counterfactual approach to causation has figured as a prominent starting point for those inclined towards a view on causation that is not as strictly reductive in spirit as Lewis's. We start out, below, by considering this counterfactual approach. We will not further discuss regularity-based theories of causation here, for this would not help us much in exploring the threshold between the first and the second conceptual gear.[244]

[242] For a fairly classical account of causation roughly along the sketched lines, see Davidson [1967], but see also the strikingly similar view defended by David Armstrong, professed anti-Humean, in his [1997, esp. chapter 14]. We will discuss Armstrong's theory below (see esp. §6.1.2). It is interesting, furthermore, that the Stanford Encyclopedia article entitled 'The Metaphysics of Causation' is organized entirely in terms of such a relation: it contains a discussion of the nature of the relation involved, of the nature of the relata involved, of the number of relata, etc. See Schaffer [2014].

[243] Recall our observations to similar effect in the previous chapter, while discussing laws of nature (see p. 171 in §5.2.2), and in chapter 3, while discussing the Tarskian, reductive approach to logical validity (see §3.2, p. 92).

[244] But see Stroud [2011, ch. 2] for a thorough and insightful critical discussion of such theories.

6.1.1 The Counterfactual Analysis

In his landmark paper on causation [1973a], Lewis starts out by quoting the following passage from Hume:

> [W]e may define a cause to be *an object, followed by another, and where all the objects, similar to the first, are followed by objects similar to the second.* Or in other words, *where, if the first object had not been, the second never had existed.* [Hume 2007/1748, VII.2, p. 56]

The first part of this quote corresponds to the regularity view on causation, but the second part forms the basis for Lewis's proposal: a counterfactual theory of causation. Lewis defines a relation of *causal dependence* between two actual events c and e which holds just if it is true that, were c not to occur, e would not occur. Causation is then defined to be the ancestral of this relation: c causes e just if there is a chain of events from c to e such that each event (except c itself) causally depends on the previous one in the chain.

Of course, this analysis depends on a further story as to how we should understand the counterfactuals. And it is here that we find the reason why Humeans, who are usually so adverse of modal notions, came to be interested in a counterfactual analysis like Lewis's at all: counterfactuals, it seemed, would need to be accounted for anyway, given their ubiquity even in scientific discourse. Here is Nelson Goodman, who played a significant role in putting the issue of counterfactuals high on the philosophical agenda of the mid-twentieth century:

> The analysis of counterfactual conditionals is no fussy little grammatical exercise. Indeed, if we lack the means for interpreting counterfactual conditionals, we can hardly claim to have any adequate philosophy of science. ... [A] solution to the problem of counterfactuals would give us the answer to critical questions about law, confirmation, and the meaning of potentiality. [Goodman 1955, p. 3]

So why not use them also for the analysis of causation? With the advent of possible worlds semantics, it seemed that decent logical handles were now available for a satisfactory theory of counterfactuals, i.e., a theory that allowed them to perform the kind of work Goodman and others required them to do. And indeed, quite early in Lewis's philosophical career, around the same time at which he developed his ideas on causation, we find his famous possible worlds-based account of counterfactuals (see Lewis [1973b]). Put very briefly, Lewis's view on counterfactuals is as follows. A counterfactual is true just if its consequent is true in the closest world(s) in which the antecedent is true, where closeness of worlds, in turn, is spelled out in terms of

qualitative similarity and comparative obedience of the actual laws of nature (the number and size of divergences from those laws, i.e., of 'miracles').[245]

There has been extensive debate on the merits and problems of this approach to causation—most notably on questions concerning the closeness ordering of worlds, the transitivity of causation, time-asymmetry, and preemption.[246] As a result, many bells and whistles have been added to the theory; most prominently, of course, by Lewis himself.[247] For our purposes, these problems and ensuing potential remedies are of minor importance. What interests us is, rather, what made this counterfactual approach so attractive.

The advantage of a counterfactual approach to causation, as opposed to a regularity view, is that it stays closer to the intuitive notion. For a change, the modality involved in the concept of causation is not subjected to the all-too-familiar strategy of reduction in terms of generality. We just quoted Goodman's statement that without 'means for interpreting counterfactual conditionals, we can hardly claim to have any adequate philosophy of science'—apparently, the modality involved is too important to ignore.

For first-gear metaphysicians, it is clear that Lewis's reductive account of modality helped make the modality used in the counterfactual analysis acceptable and tractable. By the same token, however, the advantage we just formulated disappears: the subsequent reductive analysis of counterfactuals makes short work of the modality involved. And, indeed, here generality plays a crucial role again (in *all* closest worlds which satisfy . . . , it is the case that . . .). Given the full machinery of Lewisian metaphysics, this generality affects the resulting conception of causation in rather surprising ways. For consider its reliance on laws of nature: these play a role in generating the weak ordering on worlds required for finding the closest world(s). In and of itself, this seems to be not very problematic, but, as Lewis himself observes, writing out what this means in terms of his own Best Systems account of laws spells trouble:

> Like any regularity theory, the best-system analysis says that laws hold in virtue of patterns spread over all of space and time. If laws underlie causation, that means that we are wrong if we think, for instance, that the causal roles of my brain states

[245] In his [1979b], Lewis responded to a critical notice by Kit Fine [1975] with a more precise statement of the role similarity and the laws play in determining closeness, where he opts for the following order of priority: (1) avoid major violations of natural law, (2) maximize overall exact similarity, (3) avoid small and local violations of law, and (4) maximize overall approximate similarity. See Lewis [1979b, pp. 47–8].

[246] See, e.g., Fine [1975], Jackson [1977], and Elga [2001], and the essays in Collins *et al.* [2004] and Hoerl *et al.* [2012].

[247] See the essays and postscripts in Lewis [1986b, parts 4 and 6], and the quite substantial revision of his counterfactual theory in terms of influence rather than causal dependence in [Lewis 2000, 2004a,b].

> here and now are an entirely local matter. That's an unpleasant surprise, but I'm prepared to bite the bullet. [Lewis 1994, p. 479]

In fact, things are worse than Lewis here suggests: whether or not my brain states cause something depends not only on what the actual mosaic looks like in spatiotemporally utterly remote areas, but also on what many other worlds are like—an unpleasant surprise, indeed. On the other hand, it should not be really surprising: by systematically replacing modality by generality, which is what first-gear thinkers do, one is bound to end up with distressingly extrinsic simulations of the original notions.

Be that as it may, the appeal of a counterfactual analysis of causation lies precisely in its being more faithful to the modal and productive aspects intuitively associated with causation than regularity views—witness Lewis's use of labels like 'dependence' and 'influence'. Perhaps, then, this could serve as the starting point for a conception of causation that is more faithful to its second-gear nature. Of course, in order to suit that purpose, the reductive analysis of counterfactuals will have to be rejected. Such a rejection also leaves room for a truly local understanding of causation.

This hope, we claim, underlies the currently popular manipulability approach to causation, which is explicitly not meant to yield a reductive account. The basic thought underlying this approach is that one can manipulate effects by manipulating their causes. This thought has an eminently counterfactual aspect: if someone *were to* manipulate the cause in certain ways, this *would* result in systematic variations in the effect.[248] The step from manipulation understood in terms of human agency, on which this basic thought rests, to the less anthropocentric notion of an *intervention* is then easily made: in cases of causation, there is always a possible intervention on the cause such that the effect would have differed accordingly and purely in virtue of the intervention. Or, in other words, there is always an exploitable *invariance* between cause and effect.

Now, however much illumination the philosophical work on these topics has brought for our practical understanding of causation: as far as the metaphysics of causation is concerned, the intuitively appealing use of terminology like 'manipulability' and 'intervention' is in the end mere decoration of the underlying counterfactual analysis. In fact, the decoration even gets the theory into trouble, because the requirement of a possible intervention is in certain circumstances simply too strong. E.g., the Big Bang caused many things, yet it is unclear at best whether it

[248] One of the founding fathers of this view is von Wright [1971]; the idea has been developed more recently by Menzies and Price [1993], Pearl [2000], Woodward [1997, 2003], Woodward and Hitchcock [2003], and Hitchcock and Woodward [2003].

is possible for something to have intervened on it.[249] In the end, then, manipulability approaches to causation propose that we consider counterfactuals that differ from the ones Lewis uses in his analysis, but they leave us entirely in the dark as to how we should understand their relation to causation. What we need, then, is not simply to abstain from endorsing a reductive strategy in terms of first-gear notions, but rather a positive metaphysical account of causation that *explains* the truth of the relevant counterfactuals.[250] Such an account is usually not provided by the defenders of manipulability theories such as interventionism: their aim is not to capture the metaphysics of causation but rather to clarify the position of the concept of causation within a wider circle of interrelated concepts. Hence they think they can adopt a neutral stance on the metaphysical side of things.[251] Woodward, for instance, writes:

> [C]oncepts like "cause" and "explanation" belong to an interrelated family or circle of concepts that also includes notions like "law," "physical possibility," and other modally committed notions. (The circle may also include "counterfactual dependence," depending on how this notion is understood.) . . .
> The account that I present is not reductive . . . . However, . . . I would be delighted if someone were able to show how the nonreductive characterizations of cause and explanation that I provide might be replaced by reductive characterizations. . . .
> [A]lthough I offer a nonreductive treatment of the kinds of counterfactuals that are relevant to elucidating "cause," "law," and "explanation," I also argue that the counterfactuals on which the philosophical tradition has tended to focus in elucidating these notions are the wrong counterfactuals for this purpose. Again, the correctness of this claim seems completely independent of questions of reduction. [Woodward 2003, p. 20–1]

The 'interrelated family or circle of concepts' is, of course, the second conceptual gear, and insofar we agree with Woodward that to clarify the position of causation within this circle is an interesting and important project. We do not agree, however, with his separation of this task from the metaphysics of causation—recall that, on

[249] Considerations such as these have motivated manipulability theorists to claim that the required intervention need not be *physically* possible but only *logically* possible. Logical possibility, however, leaves too much open to be of use in the analysis of causation. This criticism has been quite convincingly developed by Reutlinger [2012].

[250] This seems to be a problem attaching to counterfactual accounts in general: if the reductive approach is abandoned, there is no reason to think that the counterfactuals can play any interesting role anymore, except for illustration. (And, we may add, similar remarks apply to possible worlds in general.) The question then becomes what it is about the subject matter in question that manifests itself in such counterfactual truths.

[251] Not all defenders of the kind of approach we are discussing are thus clear on what the metaphysical status of their views is. Halpern and Pearl, for instance, think that they 'have shown that the counterfactual approach to causation, in the tradition of Hume and Lewis, need not be abandoned' [Halpern and Pearl 2005, pp. 877–8], yet their account of the underlying counterfactuals in effect depends on 'causal knowledge'.

our view, conceptual truths *are* metaphysical truths (see §3.4). Does that mean that we can turn Woodward's account of causation into a metaphysical account simply by disagreeing with him on this separation? No: by distinguishing thus sharply between the conceptual and the metaphysical side of things, Woodward relieves himself of the task of showing which are the *fundamental* concepts (e.g., causation), and how they relate to the other ones (e.g., counterfactual dependence, intervention).

Let us, therefore, continue to look for a suitable metaphysical account of causation that explains the truth of the relevant counterfactuals. It looks like what we need is precisely what the Humeans do not like: necessary connections.

6.1.2 Necessary Connections Revisited

What underlies the Humean approach to causation, both in its regularity-based guise and in its counterfactual guise is, of course, the idea of a Humean mosaic that is devoid of necessary connections. All the events are already in there; questions of causation then merely concern how to fix the extension of a specific relation on these events. Hence the Humean starts looking for suitable features of the mosaic (and of suitably similar mosaics, i.e., the closest possible worlds) on which this relation may be said to supervene—which explains the noticed tendency to arrive at a highly extrinsic explication of causation.

The advent of anti-Humean thought over the past decades has gone hand in hand with a gradual disentangling from first-gear views. We have just seen an interesting instance of this phenomenon: the counterfactual analysis of causation. It was developed within a first-gear setting, yet also provided, at least at first sight, handles for those searching for a more robust, second-gear understanding of causation. We learned, however, that it is ill-advised to distance oneself from the Humean, reductive orthodoxy by *refraining* from metaphysical commitment.

A straightforward way of committing oneself to an anti-Humean metaphysical picture would be to insist on precisely the kind of thing that the Humean is skeptical about—necessary connections. At this point, it is helpful to briefly return to laws of nature. In the 1970s, David Armstrong and others proposed an alternative to prevailing Humean views on laws of nature, in terms of *universals*: if it is a law that *F*s are *G*, they claimed, this is not simply because all *F*s happen to be *G*, but rather because there is a necessitating link between *F*ness and *G*ness—in Armstrong's terms, $N(F, G)$ holds.[252] This necessitating link, which Armstrong likes to call 'nomic

[252] See Dretske [1977], Tooley [1977], and Armstrong [1978b, 1983, 1997].

necessity', *explains* the corresponding regularity: the direction of explanation is no longer reversed. That is, the laws govern what happens in the world, and are not mere summaries of what happens in the world (see §5.2.2). As we will see in a moment, Armstrong uses his view on laws of nature to account for causality in roughly the following way: causation is the subsequent instantiation of universals which are themselves related by his relation of nomic necessity.

It is interesting to look at Lewis's reaction to this proposal concerning laws:

> Whatever N may be, I cannot see how it could be absolutely impossible to have $N(F, G)$ and Fa without Ga. (Unless N just is constant conjunction, or constant conjunction plus something else, in which case Armstrong's theory turns into a form of the regularity theory he rejects.) The mystery is somewhat hidden by Armstrong's terminology. He uses 'necessitates' as a name for the lawmaking universal N; and who would be surprised to hear that if F 'necessitates' G and a has F, then a must have G? But I say that N deserves the name of 'necessitation' only if, somehow, it really can enter into the requisite necessary connections. It can't enter into them just by bearing a name, any more than one can have mighty biceps just by being called 'Armstrong'. [Lewis 1983, p. 366]

Lewis, being a first-gear thinker, cannot but conceive of Armstrong's proposal as he does here. Like particulars, universals are, for Lewis, completely loose and separate existences (if they exist at all).[253] Hence, if N relates F and G, then this is just one 'particular matter of fact'; likewise for a's being F. And these matters of fact are entirely independent from a's being G or not. Thus, the only construal of N that would be acceptable for Lewis is either as encoding *de facto* constant conjunction of F and G (as he writes), or as conveying that G is part of F (in some sense), in which case it is trivial that Fs are always G. That is, unless content can be given to Armstrong's terminology *in first-gear terms*, Lewis will remain skeptical about that content.

In reply to just such criticisms, Armstrong sets out to explain how it is we may come to endorse such necessary connections between universals:

> It begins from 'what happens and what things are like'—the four-dimensional scenery, as I take it. This scenery exhibits many regularities .... In particular, I would argue, it includes many regularities whose instances are causal sequences. ... So, I claim, we can start from singular causations that nevertheless exhibit a pattern, a pattern where, putting it very roughly, the same type of cause produces the same type of effect. ... We have the bunch of singular causations, the same sort of cause bringing about the same sort of effect. May we not seek to explain this? May we not hypothesize that this uniformity [of Fs causing Gs] holds *because* something's being F *brings it about* that that same something becomes G? [Armstrong 1993, p. 421]

[253] In fact, Lewis and Armstrong are largely in agreement here—see below.

Here, we witness Armstrong struggling to incorporate second-gear ideas ('produces', 'brings it about') into a basically first-gear setting ('the four-dimensional scenery'—we will fill in some further relevant details of Armstrong's position shortly). What we are getting at is that insisting on second-gear ideas within a first-gear setting, as Armstrong does, is quite hopeless. In particular, a second-gear conception of causation, which is what Armstrong seems to be relying on here, involves the thought that the cause really and seriously *produces* the effect, and hence requires a radically different setting, as we will see below (§6.1.3).

Now, Armstrong nevertheless tries to hold on to this anti-Humean idea of causation. His strategy is twofold: on the one hand, he insists that singular causation is 'conceptually primitive', while on the other hand, he refers us to paradigm cases of direct perception of causation. We discuss these two parts of his view on causation in turn.

Armstrong characterizes his commitment to a conceptually primitive notion of singular causation as follows:

> That singular causation does exist, and that the world's work is done by such causing, is the natural assumption of all those who have not fallen under the Humean spell. [Armstrong 1997, p. 204]
> The concept of singular causation is ... *conceptually* primitive. ... At the same time, ... it is not ontologically primitive.[254] It can be given an ontological analysis ... as the instantiation of a law of nature. [ibid, p. 202]
> When we experience singular causation, what we are experiencing is nomicity, law-instantiation. [ibid, p. 227]

One can easily imagine Lewis being as skeptical about such a notion of 'singular causation' as he was about Armstrong's necessitating links between universals. By declaring it to be 'primitive', Armstrong refuses what Lewis demands: to provide a story as to the content of this notion. Moreover, Armstrong thus makes it nearly impossible to tell the metaphysical difference between instantiations of Armstrongian laws and instantiations of Humean regularity-based laws. After all, Lewis could also call the instantiations of his Best Systems laws cases of 'singular causation', if he wished.

But Armstrong is, of course, well aware of such worries, which is why he also frequently refers us to the fact that we can directly perceive singular causation:

> Some philosophers ... may object to my using the existence of singular causation .... But it seems clear that we (and animals) do regularly perceive that one thing

[254] As an aside, notice how Armstrong here separates ontological matters from conceptual matters, just like Woodward did. As we have argued in chapter 1, this sort of separation is typical for metaphysical realists working under the Modern picture.

> causes another. ... A conspicuous case is perception of pressure on our body. [Armstrong 1993, p. 421]
> We can be aware ... that we have successfully acted in a certain situation, that what we did sprang from our will as cause. [Armstrong 2010, p. 45]
> All we have in our concept of singular causation ... is that something ... *makes something happen*. At the conceptual level this making something happen, this singular causation, is a primitive. [Armstrong 1997, p. 210–11]

In his attempt to motivate a *truly* anti-Humean understanding of causation, and not just one by name, Armstrong here clearly moves towards second-gear thinking: we don't look anymore at the make-up of the entire Humean mosaic (the 'four-dimensional scenery') or at putative necessitating links between universals, but rather at something we can experience quite concretely: pressure on our body, and the efficacy of our own agency. Armstrong seems to be quite fond especially of the first example.[255] The examples are well-chosen: they are taken from our direct experience of the reality of causal production in perception and action. Armstrong thus clearly points to the notion of causation as genuine production that we apply in our (second-gear) dealings with the world. However, that notion stands in stark contrast to his overall view, on which causation is still considered a relation on events in the four-dimensional scenery. And Armstrong leaves no room for doubt about his stance on this point—he writes:

> Causation is a relation, fundamentally a dyadic relation, and... no relation exists without its full complement of terms. [ibid, p. 75]

We now have a more or less complete overview over Armstrong's attempt at an anti-Humean understanding of causation. It involves reference to 'pressure on our body' as a paradigmatic case of causal production, the claim that singular causation is conceptually primitive although it can metaphysically be accounted for in terms of law-instantiation, and the traditional idea that causation is a relation on events in space-time. Given these ingredients, it is not surprising that Armstrong resolutely rejects the counterfactual analysis of causation, and recognizes that the counterfactuals used in such an analysis do not exhaust the notion of causation but rather arise

[255] The example of pressure on our body refers us back to the topics of Armstrong's earliest work, especially Armstrong [1961, 1962, esp. chapter 3]. It recurs widely throughout his work—compare the following quotes: '[W]e do seem to have direct knowledge of such states of affairs as physical pressure on our body' [Armstrong 1973, p. 20]; 'The dyadic predicate "causes" is as much an observational predicate as any other predicate in our language, especially in such cases as our awareness of pressure on our own body' [Armstrong 1997, p. 228]; 'The philosophically little discussed perception of pressure on our body ... is as directly perceived, I maintain, as anything else in our experience' [Armstrong 2004b, p. 128]; and: '[I]f causation is singular, then there seems to be an obvious candidate for perception: the action of various forces upon our body' [Armstrong 2010, p. 45].

from that notion—but only in favorable cases where no preemption or prevention or intervention is lurking in the neighborhood. He writes:

> Suppose it is true that event *c* causes event *e*. We know that, because of the possibility of a backup cause, it is false that it is entailed that if *c* had not occurred, then *e* would not have occurred. ... So what entailment can we find? It is not at all clear that we can do anything better than this: "*c* causes *e*" entails that: "if not-*c*, then, in the absence of a backup cause, ... not-*e*." It may well be, of course, ... that a backup cause is absent .... If so, then the stronger "if not-*c*, then not-*e*" will be *true*. Because of this, it seems fair enough to say that the truth that *c* causes *e supports* the proposition "if not-*c*, then not-*e*." It seems to me that this throws out an important challenge to those, such as Lewis, who hold that the counterfactual analysis of causation is conceptually true. The onus is on them to give us an entailment of "*c* causes *e*" that involves counterfactuals but makes no mention of causation. I have no proof that the onus cannot be discharged, but I do suspect that it cannot be. [Armstrong 2004a, pp. 451–2]

The Humean insists that the onus *has* to be discharged somehow, most probably by moving towards ever more sophisticated analyses, or perhaps by biting some hard bullets, while the anti-Humean agrees with Armstrong here 'that it cannot be'.[256] A typical first-gear/second-gear clash.

Now, earlier we noticed that such a stance requires some story as to what makes true the relevant counterfactuals. Armstrong recognizes this requirement, and hints at what kind of story his metaphysical picture allows him to construct:

> These ... counterfactuals will have to have truthmakers, of course. Counterfactuals cannot hang on ontological air. A key part of the truthmaker will be the *relevant laws of causal action*, together presumably with the boundary conditions .... [ibid, p. 449]

There is a lot to be said on whether or not this gesture towards 'truthmakers' for causal counterfactuals will do, but one thing is clear: Armstrong attempts to build into his picture a truly anti-Humean understanding of cause and law, and uses that understanding to shed light on how and why causation supports counterfactuals. That is precisely the sort of work we criticized the manipulability theorists for not doing. Thus, we have made progress.

But the progress is not sufficient yet. We noticed along the way that Armstrong's overall metaphysical position is much more first-gear than second-gear in spirit, despite his opposition to Humeanism. He insists that the only way to get rid of

[256] Armstrong writes, while discussing prevention with the help of a type of counterfactuals he borrows from Dowe [2000]: 'Causality—causal *action*—appears as an *unanalyzed* notion in these counterfactuals. In the case of prevention, the counterfactuals give us in the consequent condition the *causal* consequences that would flow from the absence of the successful prevention.' [Armstrong 2004a, p. 449].

Humean reductionism with regard to causation is by rejecting any attempt at reductive analysis, be it in terms of regularities or in terms of counterfactuals, and hence endorses singular causation understood as law-instantiation. That, however, does not provide him with genuine production, his use of robust, second-gear terminology notwithstanding. To have true production, Armstrong would also have to give up his understanding of causation as a dyadic relation on events in the 'four-dimensional scenery': if the cause truly *produces* the effect, then causation cannot be such a relation, simply because the causing *explains* the existence of the effect, whereas a relation *presupposes* the existence of its relata. Which is not to deny that one can, of course, define a relation of 'causation' that holds between two events (or whatever the relata are thought to be) whenever the first *has produced* the second (notice, by the way, the perfective aspect of which this definition makes use). That would, however, leave the relevant notion of production entirely in the dark. Causation as production is prior to causation as a relation between cause and effect.

We learn, then, that simply adding necessary connections to a Humean mosaic, which is more or less what Armstrong does[257], does not get us outside of the realm of purely first-gear thought. That is why Armstrong's brave attempts merely invite the Lewisian response that he is in effect not departing at all from Humeanism despite using anti-Humean rhetoric.

We started out by looking at varieties of the counterfactual approach to causation. This approach, we saw, is based on incorporating aspects of the second-gear concept of causation-as-production into a first-gear setting—that is, by incorporating the counterfactuals to which genuine production gives rise. We have now seen that it is just the habit of sticking to first-gear thinking which prevents this notion of causal production from being endorsed in its full metaphysical robustness. Hence, our next step is to leave behind such habits and look for a metaphysical setting within which causation as production makes sense.

6.1.3 Causation as Production

By following one thread of philosophical debate over causation, which started from a first-gear setting, we have arrived at a second-gear concept of singular causation as production. We did so by gradually discarding the first-gear presuppositions that stood in our way: the idea of a Humean mosaic to which causation should somehow be reducible, the focus on generality as underlying causality (be it in terms of classical

[257] '[T]he great causal net of the world', as he writes—see Armstrong [2004a, p. 448].

regularities or in terms of a regularity-based reductive analysis of counterfactuals like Lewis's), and the presupposition that causation must be a relation of some sort on a domain of events.

What we arrived at up until now has already been observed some time ago by Elisabeth Anscombe in her famous Inaugural Lecture at Cambridge. She remarks:

> [C]ausality consists in the derivativeness of an effect from its causes. This is the core, the common feature, of causality in its various forms. Effects derive from, arise out of, come of, their causes .... [A]nalysis in terms of necessity or universality does not tell us of this derivedness of the effect; rather it forgets about that. [Anscombe 1971, p. 136]
> The most neglected of the key topics in this subject are: interference and prevention. [ibid, p. 137]

What Anscombe here calls the 'derivativeness' of an effect from its causes is what we have been calling causal production. We have seen that, indeed, both the Humeans and those who, like Armstrong, try to incorporate a more robust notion of causation based on necessitation into a more-or-less first-gear metaphysical setting, 'forget about' this productive aspect, or, rather, inevitably fail to make room for it.

Anscombe also provides us with a hint: that we should try to understand 'interference and prevention' if we wish to understand causation. Interestingly, the ongoing discussion on counterfactual analyses of causation strikingly supports her point: virtually all of the ever more baroque counterexamples that are being discussed involve interference, prevention, preemption, etc.[258] Understood from this perspective, the literature has brought to light very clearly that an understanding of causation as a relation on events is bound to be incorrect: typically, it is possible for any causal process to be interrupted *before* the effect it is said to give rise to occurs. We have discussed this issue to some extent in our introductory section on second-gear forms of thought (see §4.3): because of possible interference, virtually nothing can be said to be brought about 'with causal necessity', except for the causal processes themselves. Such is the *ceteris paribus* nature of most second-gear phenomena.

Typically, defenders of Humean approaches take issues like prevention and interference to be rather inessential, annoying details to be accounted for by fine-tuning their view. Armstrong, on the other hand, likes to use the possibility of interference and prevention as an argument against Humeans: singular causation does not imply any corresponding regularity, he thinks, and therefore Humean analyses fail. Yet his own view faces similar problems. We saw that singular causation just

[258] We provided some relevant references in fn. 246 on p. 186 above.

is law-instantiation, on his view, and laws are states of affairs of the form $N(F, G)$. On Armstrong's view of universals, which states that universals are wholly present wherever they are instantiated[259], we have to conclude that wherever F is present, $N(F, G)$ is present, and where $N(F, G)$ is present, G is present. So there seems to be no room for interferences on his own view either.[260]

Hence, Armstrong's view needs fine-tuning too. We could go roughly the same way as the Humeans, and build into the F-part of a putative Armstrongian law $N(F, G)$ an exclusion of everything that *could* be preventing or interfering with the causal process. In fact, Armstrong himself considered something like this in earlier versions of his theory:

> But cannot an oaken law [i.e., a law which admits interference] always be represented, in principle at least, as an iron law [i.e., a law which does not admit interference] by putting in all the negative qualifications? Yes, in a way it can, provided that we bear in mind how wide the qualifications may be which are implied by the phrase 'in principle'. [Armstrong 1983, p. 149]
> [W]hen we consider all ... potential interferences, ... [we] cannot stop short of the whole state of the universe at that time. Only then do we reach the total cause. [Armstrong 1978b, p. 155]

This move, however, creates a tension with his insistence on singular causation as a local feature of reality, as one that can even be perceived, like in the case of 'pressure on our body'. Another option, proposed by Schrenk [2011, p. 588], is to keep F as it is, and instead change G into a potential or disposition: the necessitation would then hold not between F and G itself but rather between F and a potential to become G (a 'causal factor'). The potential, of course, may or may not actualize, hence allowing for interference to occur.[261] Notice how this move in effect gives up on causation as a relation between causes and effects in favor of an understanding in terms of causal processes (which correspond, in our terminology from §4.3, to the mentioned causal factors.)

This brings us to the more robust anti-Humean views on causation in terms of dispositions or powers that have gained influence over recent decades, and thereby to the kind of picture we have sketched in some detail in our survey of second-gear forms of thought (see §4.3): here we find a truly local, i.e., non-general understanding of

[259] See, e.g., Armstrong [1997, chs. 3–4].
[260] This point has been made, quite convincingly, by Schrenk [2011], who goes on to suggest a view that moves the Armstrongian perspective a good deal closer to full-blooded second-gear thought in our sense (although not quite all the way).
[261] Schrenk reports that Armstrong sympathized with this suggestion in personal correspondence, although 'what the nature of these factors is remains to be seen'. See Schrenk [2011, p. 589].

causal production that does not depend on necessitation between causes and effects, and does not view causation as a (first-gear) relation between given events.

6.1.4 Adding Powers

In analytic philosophy, almost everything starts with Humeanism. So too in the case of powers or dispositions. At first glance, it seemed that the Humeans could make sense of such modal properties in terms of conditionals: to say, for instance, that a glass is fragile, i.e., that it has a disposition to break, is just to say that it breaks if struck. Obviously, this crude analysis in terms of a material conditional only works for objects that are in fact struck at some (presumably late) point of their existence, for otherwise it would turn every object that is never struck into a fragile object. Hence something more is needed: the glass *would* break if it *were* struck, even if it is never actually struck. In other words: a counterfactual analysis[262]—and there we are, back where we started.

Luckily, at this point of our argument we can be rather brief, because the brand of anti-Humeanism we are now interested in, that of the dispositionalists, rejects this approach straight out of hand.[263] The reason should by now be obvious: the counterfactuals that the Humean wishes to employ for his analysis of dispositions are true (if at all) *because of* the dispositions, and not *vice versa*. It is interesting to note that, just like in the case of the counterfactual analysis of causation, recent decades have yielded a huge literature on increasingly complicated counterexamples, in this case involving so-called masks, antidotes, finks, mimickers and the like.[264] And, as before, all such counterexamples to the ever more sophisticated counterfactual analyses of dispositions vanish once we adopt the mentioned change of explanatory direction: the relevant counterfactuals may fail to hold in specific cases without that having anything to do with whether or not the disposition in question is present. Indeed, the elaborate stories in which such counterexamples occur often themselves contain the very reason why the counterfactual fails to hold (e.g., the glass is fragile,

[262] This analysis has been defended, e.g., by Ryle [1949, esp. ch. V], Goodman [1955, ch. 2], Sellars [1958], and Quine [1960, §46].

[263] Or, at least, such anti-Humeans *should* thus reject such a counterfactual analysis. It is surprising to what extent those writing on dispositions or powers, despite not committing themselves to Humeanism, insist on being bothered by counterexamples to the counterfactual analysis.

[264] On masks and antidotes, which involve special situations in which a certain disposition fails to manifest even though it is present, see Johnston [1992], Bird [1998], and Choi [2003]. On finks, which involve situations in which a disposition is taken away by the very same circumstances which are thought to trigger its manifestation, see Martin [1994]. On mimicking, which involves situations that make true the relevant counterfactuals even though the disposition is *not* present, see Lewis [1997]. That such 'exceptional' situations are in fact quite ubiquitous is argued by Fara [2005].

but it is packed in styrofoam, hence it will not break when struck).

The close connection between dispositions or powers on the one hand and causation on the other thus shows up even on this level: the analysis of the former invites precisely the same sort of problems as we saw in the case of analyses of causation. Hence anti-Humeans aim to incorporate dispositions or powers as basic metaphysical ingredients into their overall metaphysical view, in order to use them for their understanding of causation. They are placed, of course, within the realm of properties. Some writers, the 'pandispositionalists', do so by declaring that *all* properties are, in the end, powers[265]; others say that *some* of the properties there are, are powers, thus allowing for merely categorical properties as well[266]; still others claim that every property has both a categorical and a dispositional aspect or nature.[267] We do not intend to engage in a discussion on which of these three alternative views is to be preferred, even though the question as to the nature of properties is an interesting one also for our project of understanding the differences between our three conceptual gears. Suffice it to say that all three options evidently make use of second-gear forms of thought—the distinction between state-predication and process-predication we introduced in our survey of second-gear thought (see §4.3) mirrors the distinction between categorical and dispositional properties. The three views we have just named can then be framed as claiming, respectively, (1) that, contrary to our claim, the second conceptual gear does not allow for state-predication; (2) that state-predications and process-predications are at least sometimes true independently from one another; and (3) that all state-predication implies process-predication and *vice versa*.

If we go along with the dispositionalists, the metaphysical picture we end up with comes very close to the sketch of the second conceptual gear we provided in §4.3: it is the true second-gear picture of physical objects that manifest their various causal powers under certain conditions, where these manifestations consist in the occurrence of corresponding causal processes (Schrenk's 'causal factors'). Causal

[265] See, e.g., Mumford [2004], Bird [2005, 2007], Chakravartty [2007], Whittle [2008], and Marmodoro [2009]. Hüttemann [2013] is not very explicit about his views on properties but clearly sympathizes with a powers-only view. See also Lowe [2006, p. 138] and Oderberg [2012] for critical discussion. It is not always clear whether defenders of pandispositionalism intend their claim to extend merely to the realm of the properties, such that their ontology includes both powers and their bearers, or rather so far as to entail that *all there is* is powers, such that even the bearers of powers are in the end themselves (complexes of) powers. Perhaps the latter view can be understood to be the dispositionalist's version of a pure process ontology (see fn. 189 on p. 137).

[266] See, e.g., Swoyer [1982], B. D. Ellis and Lierse [1994], B. D. Ellis [2001, 2010], and Molnar [2003].

[267] See, e.g., Martin [1997], Martin and Heil [1999], Heil [2003, 2005], Mumford [1994, 1998], and Mumford and Anjum [2011].

processes unfold in certain ways but can be interrupted or interfered with by other processes. At best, then, we can say that an object which has a certain causal power will, when put in the right conditions, *ceteris paribus* produce the effect to which its manifestation, the corresponding causal process, gives rise. That is, no such effect distinct from the causal process itself will *of necessity* or even *invariably* be produced. Take, for instance, the power of ^{235}U to indulge in a chain reaction of nuclear fission, on which nuclear reactors are based: in fact, all of our reactors use control rods to make sure the chain reaction does not run out of hand, and hence there are virtually no actual cases in which the process that constitutes the manifestation of that power unfolds without interference.

The idea of powers or dispositions as part of the fundamental make-up of reality has long been the target of strong skepticism or even ridicule, as in the classical case of Molière's play *Le malade imaginaire*: to explain something's producing X (say, some opium's putting its consumer to sleep) by ascribing to it a power to X (a 'virtus dormitiva') seems rather uninformative. And indeed it is—but only as long as we refrain from further investigation. If it is possible to get a grip on the precise conditions under which the power is manifested as well as on the ways in which the process leading to X unfolds and interacts with other processes, the charge of uninformativeness disappears—and without recourse to reductive analyses. On our view, to state the second-gear generic thought that opium has a power to produce sleepiness in humans *is* to state that there are such conditions, that there is such a default way for the process to unfold, and that there are ways in which it interacts with other processes. Those are all conceptual truths, after all, even if we do not yet know them. And, in the case of opium's 'virtus dormitiva', it is even possible to investigate what it is about opium and what it is about humans that makes it the case that opium has the effect it has. It is possible, that is, to discover the 'mechanism' by which opium produces sleepiness (we touch upon this notion of mechanism in §7.2). But all of this starts with the observation that, apparently, opium *has* the relevant power. Appearances may be misleading, of course—for all we knew at first sight, things could be different, say, in that people with a strong propensity to develop sleepiness tend to consume opium just before that propensity kicks in. That is why we need scientific experimentation—or just everyday methods of checking—in order to find out what is really so. Notice that what we thus find out is the formal aspect of causation, as we described it in §4.3, p. 139: the essences of the relevant objects and powers determine what *can* happen, by determining what processes will start in response to which circumstances. The efficient cause realizes one of those causal

possibilities by bringing about the relevant circumstances.

We criticized Armstrong's insistence on a primitive concept of singular causation on the grounds that its primitiveness does not endow it with the desired content of causation-as-production. Are we not susceptible to a similar criticism in maintaining that powers and their corresponding causal processes are 'primitive concepts'? We claim that we give content to these 'primitive concepts' by showing how they form part of a larger conceptual nexus, complete with its typical forms of predication and the kinds of inferences it allows for—the second conceptual gear. But does not Armstrong also provide a comprehensive picture, where his concept of singular causation is tied, e.g., to his views on universals and their instantiation, and to his theory of laws of nature as nomologically necessitating links between universals? Well, that is precisely where the problem lies. To see more clearly what we have in mind here, it is instructive to look at Armstrong's reasons for *rejecting* dispositions or powers:

> It is not denied that statements attributing dispositions and/or powers to objects, or sorts of objects, are often true. But the truth-makers or ontological ground for such true statements must always be found in the actual, or categorical, properties of the objects involved. [Armstrong 1983, p. 9]
> [H]ow can a state of affairs of a particular's having a property enfold within itself a relation (of any sort) to a further first-order state of affairs, the manifestation, which very often does not exist? ... Properties are self-contained things, keeping themselves to themselves, not pointing beyond themselves to further effects .... [Armstrong 1997, pp. 79–80]

Armstrong holds on tight to a specific mode of thought, which is strikingly first-gear in nature: ontologically speaking, he restricts himself to a conception of existents that is close to what we have called first-gear entities (see §4.2). He is suspicious towards everything that requires more than the mere first-gear existence of entities. Even predication is, on his universals-based view, nothing but the existence of the universal 'in' a state of affairs that also contains the particular.

Moreover, he adds to this, in a truly Humean spirit, the idea that the basic states of affairs that compose reality are entirely independent from one another:

> [E]ach state of affairs is a contingent being. ... [Their] constituents are (thin) particulars and universals. The universals may be monadic, that is properties, or polyadic, that is relations. The relations are all external. Every truth ..., so runs the hypothesis, finds its truth-maker or truth-makers here. [ibid, p. 139]

Armstrong finds inspiration for this picture in Wittgenstein's *Tractatus*, which he approvingly quotes:

> There is no possible way of making an inference from the existence of one situation to the existence of another, entirely different situation. [Wittgenstein 1921, 5.135, as quoted by Armstrong on p. 140]

Interestingly, this independence also applies to states of affairs concerning universals, such as Armstrong's laws of nature. The necessity that binds cause and effect, on Armstrong's view, is merely nomic necessity, not metaphysical or logical necessity. Hence even the laws of nature are contingent—the state of affairs $N(F, G)$ might fail to exist.

Thus, Armstrong feels at home within a realm of entities of various categories—particulars, universals, states of affairs—that he treats as if they were first-order entities occurring as building blocks in a Humean mosaic. And, like Lewis's thesis of Humean supervenience, Armstrong holds that 'what supervenes is no addition to being' [Armstrong 1997, p. 12], hence whatever there is over and above all the independent states of affairs has to supervene on them, that is, has to 'find its truth-maker there'.

This first-gear setting allows only for first-gear forms of thought, and that is the reason why Armstrong's insistence on 'nomic necessity', 'singular causation', 'bringing things about' etc. can be unmasked as paying mere lip service to the second-gear ideas he tries to incorporate. That is why his insistence on primitive singular causation is not to be compared with our insistence on primitive powers and processes: by thus insisting, we insist on the need for shifting to the second conceptual gear. Armstrong's apparently second-gear approach to causation breaks down because he fails to distance himself from exclusively first-gear thought.

6.1.5 Concluding Remarks on Causation

The true, second-gear notion of causation requires a metaphysical view involving causal powers, whose manifestations are causal processes that unfold through time. Attempts to isolate this notion of causation from its embedding within such a broader, second-gear conceptual setting are futile: by doing so, one ends up with, at best, a first-gear simulation of it. We have seen that Armstrong does just this: he takes causation to be a relation on a presupposed domain of events in the 'four-dimensional scenery', yet says of that relation that it really is singular causation 'taken as a primitive', and illustrates what he has in mind with the words 'singular causation' with examples such as felt pressure on our body—examples that suggest genuine production in the second-gear sense. We have seen that this is mere insistence on something that simply does not fit into Armstrong's broader metaphysical scheme.

We have not developed anything like a full-blown second-gear theory of causation here. To do so would require spelling out in much greater detail than we have done how exactly the second conceptual gear works. By contrast, it was our aim here to defend the *need* for such a project rather than embarking upon it ourselves. We have illustrated how the distinction between the first and the second conceptual gears, as outlined in §4.2 and §4.3 respectively, relates to parts of the contemporary discussion on causation, and from that exploration we may conclude that fundamentally there are only two ways to go: fully first-gear and reductive, or fully second-gear and non-reductive. In that respect, recent developments are promising: with the rise of the dispositionalists, the way has been cleared for the development of a thoroughly second-gear understanding of causation from which every element of first-gear reductionism has been purged. First and foremost, the confusing idea of causation as a relation on a domain of events in space-time has to be abandoned.

That brings us to our next topic. We have carefully avoided the issue of time in our discussion of causation, yet it is clear that time plays a central role. We did notice that part of the problem, for Armstrong, is his assumption of a 'four-dimensional scenery' that contains all the events (or states of affairs) between which the causal relation is supposed to hold. Indeed, the very idea of causal production, with which we ended, calls for a thorough revision of this picture. It is to this task that we now turn.

6.2 Time

In our presentation of first-gear metaphysics (§5.1), we indicated that a commitment to purely first-gear forms of thought leads naturally to a conception of time consisting of eternalism, B-theory and perdurantism. It will be helpful to briefly recall how these aspects of the first-gear conception of time interrelate.

We observed that the guiding thought of a first-gear understanding of time is that times are like locations. *Eternalism*, then, is the view that time constitutes a dimension to reality that is, in key respects, like the spatial dimensions, such that reality is a spatiotemporal manifold that encompasses everything past, present and future. Eternalism provides us with the picture of a time line that consists of successive moments, each decorated with its own collection of temporal truths.

B-theory provides a conception of these temporal truths. It is the view that temporal truths are best construed in tenseless terms, i.e., by use of what we called first-gear feature predication (see §4.2). They are assigned to their proper temporal location by

way of a built-in reference to that location. For example, the tensed temporal truth 'you are reading (tensed) this chapter' is analyzed in terms of the tenseless truth 'you read (tenseless) this chapter at time t' (where t denotes the present moment). The ordering of these temporal locations can be formulated satisfactorily in terms of the earlier-later relation on times. From a metaphysical point of view, no reference to a special moment which is 'the present' is required.

Finally, *perdurantism* is the view that objects persist through time by having temporal parts at each of the times at which they exist—or, more accurately, by being spread out over multiple temporal locations in the same way they may be spread out over multiple spatial locations. Thus, perdurantism augments the B-theorist's conception of temporal truth, by providing a view on how the temporal location is built into such truths: the relevant temporal part of the object(s) of which something is predicated suffices to determine the location(s) on the time line to which a given temporal truth belongs. For instance, the tensed truth 'you are reading (tensed) this chapter' is accounted for in terms of the tenseless truth 'your temporal part at t reads (tenseless) this chapter' (where t denotes the present moment).

On the second conceptual gear, by contrast, time cannot be divorced from the second-gear form of predication, which is tensed (and aspected—see §4.3). Time is not a dimension of reality, but rather a form of being that is radically different from first-gear, atemporal being. And with that temporal form of being comes a specifically temporal form of predication (exemplification). Thus time is not accounted for in terms of atemporal feature predication which is supplemented by a reference to a certain location on the time line, but rather time is taken to *consist in* the applicability of second-gear forms of thought—in a way, then, time *is* the second-gear form of predication. Things to which this form of predication can be applied are not atemporal, first-gear entities, but rather persisting physical objects. And what can be predicated of such physical objects by using second-gear predication are not mere first-gear features, but rather states of the objects, with respect to which they can change, as well as the activities by which such change is effected. By being subject to this form of predication, by thus exemplifying their properties and activities, physical objects are persisting things capable of causal interaction in the productive sense we have discussed in the previous section (and in §4.3).

Now, one might guess that once we adopt this second-gear form of predication, we are naturally lead to the opposites of all three first-gear views, that is, to presentism, A-theory and endurantism. And, on the one hand, that is true: the second conceptual gear indeed rejects eternalism, B-theory and perdurantism, for the simple reason

that these views are built on first-gear, atemporal feature predication (and cannot coherently be formulated otherwise). On the other hand, however, it turns out that this does not automatically turn presentism, A-theory and endurantism into the correct views, from a second-gear point of view. For, as we will see, all three of these views, both when taken separately and when taken together, can be understood in such a way as to *not* involve the second-gear form of predication but rather merely first-gear feature predication—albeit in a rather complicated manner. Such first-gear versions of presentism, A-theory and endurantism are, in effect, 'non-standard' forms of first-gear metaphysics. The motivation for such views derives from a dissatisfaction with the way in which standard first-gear metaphysics, as we described it in §5.1, reduces time to merely another dimension, replaces temporal truths by atemporal truths, and reduces persistence to a mere ordering of parts. That first-gear picture provides, in the eyes of many, an unsatisfactory view on reality as a four-dimensional, completely static 'block universe', and hence they—unsuccessfully—try to resist it by endorsing said versions of presentism, A-theory, and/or endurantism. The situation is, thus, much like the one we described in the previous section on causation: there, we saw that the fundamental bias towards understanding causation as a relation on a given domain of events makes it very hard to even formulate a thoroughly second-gear view on the matter. Here, we will see that it is essentially the same picture that causes trouble: the 'given domain of events' constitutes the eternalist's time line. This seemingly natural view on temporal reality as being spread out over a temporal dimension bedevils even some of the most dedicated A-theorists/presentists/endurantists. The crux of the matter is, as said, the thought that *times are like locations*. We will discover that, wherever this thought stands in the background, the resulting view on time turns out to be, in the end, based on purely first-gear forms of thought. And, on that basis, the only things such a view can allow for are first-gear, atemporal entities, while the only properties such a view admits are first-gear, atemporal features.

We have chosen to discuss time in an unusual way. On the one hand, we are arguing against first-gear approaches to understanding time by showing how they fail to capture time as a central category of the second conceptual gear. On the other hand, we aim to clarify what the second conceptual gear involves by showing how initially promising views on time can still be covertly first-gear in character. The route we take through the complex landscape of the philosophy of time is as follows. We start with a discussion of persistence in §6.2.1. We show that at least some respectable versions of endurantism amount, in the end, to mere variations on a first-gear un-

derstanding of persistence as an object's occupying distinct temporal locations, and hence are no better off than perdurantism when it comes to the metaphysics of time. The only version of endurantism which promises to escape such worries is a version that combines endurantism with A-theory, and therefore we move on to discuss tense in §6.2.2. As it turns out, A-theory, at least on some respectable versions of that view, rests upon an eternalist conception of time as constituting a time line, and hence is no better off than B-theory when it comes to the metaphysics of time. The only version of A-theory which promises to escape such worries is a version that combines A-theory with presentism, and therefore we move on to discuss presentism in §6.2.3. Now, even here, we discover that presentism, at least on some respectable versions of that view, amounts to a mere restriction on the very same eternalist conception of time as a time line that caused trouble for A-theory, and hence is no better off than eternalism when it comes to the metaphysics of time. Only when we formulate a version of presentism with the help of the second-gear form of predication do we arrive at a conception of time that does not boil down to a first-gear simulation. Moreover, this version of presentism requires a version of A-theory that survives the criticisms we mounted before. Such a version of A-theory is our target in §6.2.4. And we thereby assemble all the materials we need for a proper version of endurantism too, as we go on to set out in §6.2.5. We conclude this section with a few brief remarks on two aspects of time we otherwise neglect: the notion of an open future, and the import that the theories of relativity have for our views on time (§6.2.6).

6.2.1 The Varieties of Endurantism

The point of disagreement amongst endurantists and perdurantists is often taken to be whether or not persisting things have temporal parts: perdurantists say they do, endurantists say they don't. Alternatively, perdurance may be characterized as persistence by way of temporal extension, while endurance amounts to lack of such temporal extension.[268] The very idea both of temporal parts and of temporal extension depends, of course, on the conception of times as locations we have identified

[268] Katherine Hawley observes that endurantists may in principle accept temporal parts on top of enduring things, and that perdurantists may in principle reject temporal parts in favor of temporally extended but (mereologically) atomic perduring things. She goes on to suggest that whether or not things have temporal extension is in turn best understood in terms of acceptance or rejection of an atemporal notion of parthood: perdurantists accept that things have parts *simpliciter* while endurantists insist that, for persisting things at least, parthood is always time-relative. See Hawley [2002, §1.6]. As an aside, we should note that Hawley defends a 'third option', which is metaphysically very similar to perdurantism: stage theory, which says that the persisting things we talk about are not full-blown four-dimensional perduring things but rather temporal slices thereof. See also Sider [2001].

to be the core tenet of a first-gear understanding of time: temporal extension just is occupation of a certain temporal region, and an object's temporal parts are those by virtue of which it occupies the temporal regions that those parts occupy. Now, the opposing endurantist notions can be understood as just a different take on occupying temporal regions. Benovsky, for instance, simply states:

> Endurantism says that objects persist through time by being wholly multiply located at different *times*. [Benovsky 2009a, p. 240][269]

Thus, the popular endurantist's slogan that things are 'wholly present' whenever they exist is perfectly consistent with the conception of times as locations: something is wholly present at t, so the thought goes, just if all of it is temporally located at t.[270]

We argue that versions of endurantism that are consistent with the times-as-locations picture are invariably based on the atemporal, first-gear form of feature predication—that is, on B-theory. Like perdurantism, such endurantisms are ways of spelling out the B-theoretic understanding of temporal truth in terms of atemporal predication coupled with an explicit time reference. To show that this is so, we consider a much-discussed argument in favor of perdurantism: the so-called argument from 'temporary intrinsics'. Temporary intrinsics are properties that an object intrinsically has (at least, *prima facie* intrinsically) but only for a certain period of its existence.[271] The argument from temporary intrinsics concerns the question what exactly the having of properties amounts to for persisting objects—that is, it concerns predication as applied to persisting things, which is why it is important for our purposes. We discuss the way in which perdurantism and certain versions of endurantism treat temporary intrinsics, thereby showing that the endurantist's mere insistence on the numerical identity of an object over time (that it is 'wholly present' whenever it exists) does not make much difference for the underlying, B-theoretic understanding of temporal truth. Thus understood, the endurantism/perdurantism debate is not a debate between first- and second-gear thinkers, but rather a purely first-gear debate.

Consider the question what it is for a banana to be first green and then yellow—where, for simplicity, we take being green and being yellow as examples of (incom-

[269] Benovsky argues, in this paper, that endurantism implies trope theory *and* the view that objects are multiply instantiable—on the basis of considerations involving time travel. He fails to note that these rather implausible consequences only follow (if at all) when his narrow, rather implausible first-gear conception of endurance is presupposed.

[270] We will see, in §6.2.5, that the phrase 'wholly present' only starts to make sense if the thought that times are like locations is discarded in its entirety. But, then, it does no longer form a proper contrast with the perdurantist's partial presence.

[271] The notion of 'temporary intrinsics', as well as the argument based on that notion, was introduced by Lewis [1986a, see esp. pp. 202–4].

patible) temporary intrinsics.[272] Typically, the perdurantist will argue that, since colors such as being green and being yellow are intrinsic properties, we have to accept their doctrine of temporal parts. For how can *the very same banana* be intrinsically green *and* yellow? Only by having one (temporal) part that is intrinsically green and a distinct (temporal) part that is intrinsically yellow. The endurantist rejects the idea that objects divide into distinct temporal parts, so what should he say?

Speaking quasi-formally, the problem of temporary intrinsics is a challenge to make sense of two statements being true which *prima facie* are of the form Gb and $\neg Gb$, without this leading to a contradiction (where we could, for the sake of concreteness, let b stand for our banana, and G for being green). The appearance of contradiction results from the absence of a time reference: the question thus is where to put that time reference.[273]

The perdurantist puts it on the object-side, thus specifying the relevant temporal part of the banana, yielding Gb_t and $\neg Gb_{t'}$. This resolves the apparent contradiction, without giving up the intrinsicality of the property G.[274]

One version of endurantism, which we call *time-relative endurantism*, resolves the appearance of contradiction by adding an argument place to the property, yielding $G(b, t)$ and $\neg G(b, t')$, which says that our banana stands in the *green-at* relation to time t but not to t'.[275] This view gives up on the intrinsicality of the property G.

A related option consists in replacing the predicate involved by a time-specific predicate, yielding $G_t b$ and $\neg G_{t'} b$, which says that the banana has the property *green-at-t* but not the property *green-at-t'*.[276] Whether or not this view amounts to giving up the intrinsicality of the property G is a tricky question.

And finally, there is a third endurantist option, which is known by the name of *adverbialism*. It states that the time-index modifies neither the object nor the property

[272] The intrinsicality of being green and being yellow is, of course, a matter of much debate—but that is largely irrelevant to our discussion. Even if *all* properties turn out to be extrinsic, in the end, then there is still a question whether their extrinsicality involves a reference to times or not—which is what the argument from temporary intrinsics is really about. This point is often overlooked, but see Eddon [2010], who shows that the argument from temporary intrinsics indeed rests on the claim that the fundamental features of things (intrinsic or not) are relations involving times, for the endurantist.

[273] An interesting discussion of the debate on persistence in just these terms can be found in Rödl [2012/2005, ch. 3, §2].

[274] See, e.g., Lewis [1986a, esp. pp. 203–5], Sider [2001, esp. pp. 1–3], and Hawley [2002, ch. 1].

[275] This relational approach has been defended by Mellor [1981, ch. 7] and van Inwagen [1990], and is considered in a sympathetic manner by Rychter [2008] and Eddon [2010]. Though we should note that Mellor abandoned the view, in later work, because it seemed to him to allow things having properties relative to times at which they do not exist. See Mellor [1998, ch. 8].

[276] Defenders of such a view, also sometimes called 'indexicalism', include van Inwagen [1990] and Correia and Rosenkranz [2012, see esp. §6]. We will discuss the latter authors' understanding of time in some detail in §6.2.4 below.

but rather the way in which the object has the property, much like, e.g., the adverb 'fast' can modify the way a certain activity (*running*) applies to a certain object (*Forrest*) without relativizing anything. The result, though difficult to present formally, would be something like b is$_t$ G and $\neg(b$ is$_{t'}$ $G)$[277]

It should be clear, from our quasi-formal presentation so far, that all of the presented formulae involve a tenseless form of predication: first-gear feature predication (on which predicate logic is, after all, based). All of the sketched views are indeed alternative ways of endorsing B-theory (although adverbialism may form an exception—see below).

The entire problem of temporary intrinsics only comes about once we start formulating things purely in terms of this atemporal form of predication. For as long as we said 'the banana *was* green but *is* now yellow', there was not even the semblance of a contradiction.[278] Yet as soon as such a statement is formulated using atemporal predication, contradiction suddenly threatens, since the tenses disappear. That is what B-theory does: it starts from atemporal, first-gear feature predication and then requires that temporality be accounted for by putting a t somewhere—on the object-side, on the predicate-side, or on the predication itself. The disagreement between endurantists and perdurantists then comes down to a disagreement *amongst first-gear metaphysicians* as to where to put that t. The structure of the debate over persistence is often presented in a way that illustrates this diagnosis—in particular, of course, by those defending perdurantism.[279]

If our diagnosis is correct, then all of the views on persistence we have listed rest on atemporal feature predication, and by that very fact imply a construal of physical objects as mere first-gear entities. Some reflection reveals that this is indeed correct.

For the perdurantist, objects are spread out in time by having temporal parts occupying the relevant temporal locations. These temporal parts instantiate their properties (and relations) atemporally—the question as to *when* yesterday's banana-

[277] The adverbalist proposal has been offered independently but (almost) simultaneously by Johnston [1987], Lowe [1987, 1988], and Haslanger [1989]. We should note, however, that most of them also endorse A-theory, whereas we discuss adverbialism, at this point in our discussion, without incorporating that commitment. For critical discussion of the adverbialist view, see, e.g., Hawley [2002, §1.5].

[278] Witness Aristotle, who formulated the law of non-contradiction as follows: 'It is impossible for the same thing *at the same time* both to be-in and not to be-in the same thing in the same respect' [Aristotle 1998, p. 88, 1005b19f, emphasis added].

[279] See, e.g., Lewis [1986a, pp. 202–3 and p. 210], Sider [2001, p. 3 and ch. 3], Hawley [2002, pp. 31–2], and Hirsch [2009, p. 231–2] (though we should note that Hirsch argues that the dispute between endurantists and perdurantists is merely verbal). Contrast those with how endurantists state their own view—e.g., Merricks [1994], Wiggins [2001, p. 31], Crisp [2003, pp. 215–7], and Fine [2006]. The former tend to talk in terms of different ways of occupying time, while the latter tend to object to the very idea of occupying time. Mellor [1998, pp. 86–7] forms an exception—but then, he defends endurantism in the context of his defense of B-theory.

slice b_t is green is simply ill-formed. As Sider writes:

> [Perdurantists] tend to employ an atemporal notion of exemplification of properties and relations. [Sider 2001, p. 56]

Temporal truths are understood in terms of this 'atemporal notion of exemplification', that is, in terms of first-gear feature predication. Reality, then, contains lots and lots of temporally thin objects that are ordered by external (spatial and temporal) relations, and that together compose various temporally extended objects—space-time worms. The mereologically atomic items have their properties atemporally, temporality enters the picture by way of their temporal interrelations. Perdurantism and B-theory are virtually made for each other.

The collection of truths about a given object, both concerning its features and its relations to other objects, is, on the perdurantist's view, just a collection of atemporal truths, very much like the collection of truths concerning, say, a given triangle on the two-dimensional Euclidean plane. True, *words* are being used that normally have a temporal meaning, but here we may actually cite one of perdurantism's most dedicated defenders, Ted Sider:

> A stationary sine wave spatially "moves" up and down; unlike a "static" straight line, its height "changes" as one progresses from left to right. [Sider 2012, p. 293]

Although intended by Sider to show that accounting for change in terms of variation amongst temporal parts is not as queer as many say it is, we can use his example in the reverse direction, to show that his perdurantist view, in effect, has no way of distinguishing *real* motion and change from such 'moving' and 'changing' as in the example.[280]

So far, so good—perdurantism is meant to be a first-gear view, after all. Turning now to time-relative endurantism, which replaces properties by relations to times such as $G(a, t)$, we observe a similar reduction. In this case, the times are explicitly added as entities of their own. The thought is that physical objects stand in relations to times *atemporally*, so that the question as to *when* the object is thus related is rejected—just as in the case of the perdurantist. In fact, it does not really make sense to accept such a question, for it would allow for the thought that some object stands in the bent-at relation to t now, but no longer stands in that relation to t tomorrow.

[280] Sider may object: the time dimension can be distinguished from the spatial dimensions, on his four-dimensionalist view, and this distinction also sets apart movement and change in the temporal direction from 'movement' and 'change' in the spatial direction. Though true, such a distinction is quite superficial: notice that rotating his stationary sine wave in the four-dimensional manifold, such that its 'movement' ends up being in the temporal dimension, suffices to turn 'movement' into movement.

Now, as Rychter writes concerning the atemporality of predication in the case at hand:

> [A]n atemporal perspective will show the banana *somehow* outside time, and bearing different relations to different times. These are the relations in virtue of which, *from a temporal perspective*, the banana has different colors at different times. [Rychter 2008, p. 165]

That is, metaphysically speaking we have atemporal entities which stand in various relations (such as the *bent-at* relation) to various 'times'. These entities are located in time in a derived way, namely, by standing in such relations to certain times. Thereby, the proponent of time-relative endurantism is in the same boat as the perdurantist: persisting objects are characterized atemporally by how they relate to various times; they 'move' and 'change' by standing in different relations to different such times. Just as the perdurantist has to insist that change *just is* variation amongst temporal parts, so the time-relative endurantist has to insist that change *just is* variation in the relations an object bears to different times.

The situation is similar for the alternative endurantist position involving time-specific properties, such as G_t and $G_{t'}$: the object turns into an entity 'outside of time', bearing various rather curious properties atemporally. Both theories are, thus, essentially atemporal. Again, B-theory stands in the background: temporal truth is construed in terms of atemporal predication augmented by a reference to the relevant position on the time line. Like perdurantism, the time reference is implicit—this time not encoded into the object but rather encoded into the property.

Turning now to the adverbialist endurantist, we observe that some credible story is required as to what it means to adverbially modify a property-ascription by a time-reference—and perhaps the view we will eventually set out (in §6.2.5 below) could be phrased as such a story. But on any interpretation that treats the property-ascriptions as *atemporal* predications which are subsequently modified by the adverbial time-references, it is clear that what we have is not unlike the time-relative version of endurantism. The theory will be stated in an entirely atemporal manner, turning objects into mere first-gear entities, and turning change into mere variation amongst such entities. B-theory stands in the background once more.

To be sure, proponents of any of the above theories can insist that their respective theories simply tell us all there is to say about persistence, and that objecting to them just because they involve a B-theoretic background assumption and a construal of objects as first-gear entities is begging the question. And that is, in a way, correct—but it also shows that we have succeeded in our argument to the effect that endurantism

and perdurantism do not differ much insofar as the metaphysics of time is concerned, at least as long as B-theory stands in the background. Thus, we may agree that, as long as one insists on restricting oneself to first-gear thought, 'time' *just is* the directed order of temporal parts or times or time-indexed properties or adverbial modifications. By our lights, that is what constructing a first-gear simulation of time consists in.

Endurantism, thus, is by itself not enough to distance oneself from first-gear thought and the corresponding simulation of temporality. What is crucial is the forms of predication that one allows: as long as these are, at root, atemporal, we are dealing with mere variations on a first-gear theme. As long as we stick to mere feature predication, time is inevitably extraneous to predication, and hence it can only enter the picture as some sort of location analogous to a spatial location—which is just what B-theory consists in.

Williamson aptly describes the situation we find ourselves in, once we take the tenseless form of predication as basic, as follows:

> [W]hat is most striking here is that something temporary, that I am sitting, is claimed to consist in something permanent, that Sit_0. ... Thus what appears as transience in the object-language is treated in the metalanguage as a mere difference between times. The pertinent relations between the given times and what happens at them are not themselves temporary. ... In this way, all temporary truths are constitutively explained by purely permanent ones. If so, change is no very radical phenomenon. In that sense, the explanatory priority of times to tense excludes *radical change*. A truth is radically transient if and only if it does not consist in permanent truths. [Williamson 2013, p. 408]

We have observed that there are different ways to deny radical transience in Williamson's sense: perdurantism is one way, and the versions of endurantism we considered are other ways.

However, our survey of endurantism so far is not complete (nor was that our intention). As one might expect, there are defenders of endurantism who insist on radical transience, and thus reject B-theory as a suitable background view for their theory of persistence. So, for instance, Merricks, who argues that there really is no 'problem of temporary intrinsics', if endurantism is understood correctly. He writes:

> Other ways of dealing with the problem [viz., the kinds of endurantist positions we have discussed so far] [are] inferior because rather than show the problem dissolves given the right understanding of endurance, they instead seem to think the problem requires a positive solution. It is as if they concede that, if there is endurance, there is the problem—but they can solve it. [Merricks 1994, p. 182]

We have found ourselves largely in agreement with Merricks here: endurantism faces the 'problem of temporary intrinsics' only if it agrees on the restriction to first-gear feature predication, that is, to B-theory. But that restriction is not mandatory, and accepting it should even be seen as a mistake on the part of endurantists.

Merricks has something to say as to what the 'right understanding' of endurantism is as well:

> Defenders of the view that one ought to "take tense seriously" hold that the difference in tense between the propositions '*b* was *G*' and '*b* is not *G*' mirrors some real difference in the world—this difference is why '*b* was *G*' and '*b* is not *G*' are not contradictory. [Merricks 1994, p. 170; example adapted]

This looks promising: Merricks in effect proposes to resist translation of the tensed statements we started with in terms of a tenseless form of predication—he defends radical transience in Williamson's sense. There is no way in which our banana *b tenselessly* has properties like being green and being yellow; all we can say is that, with the passage of time, the banana changed from being green to being yellow, through a (causal) process of ripening.

The 'real difference in the world' that we need is, thus, genuine passage of time, as opposed to a mere juxtaposition of 'times' that is close to indiscernible from a juxtaposition of places. Genuine passage is the feature of time that differentiates real change and motion from the 'change' and 'motion' involved in Sider's (stationary) sine wave (see above). Real passage is what Williamson's notion of radical transience and radical change points towards. A 'difference in tense' that is to be 'taken seriously' in roughly this way sounds, thus, very much like our second-gear form of predication.[281]

It looks, then, as if a proper understanding of tense will allow us to formulate a version of endurantism that does not rely, explicitly or implicitly, on B-theory, with its atemporal understanding of objects, properties, and predication. What this tensed version of endurantism exactly looks like will be the topic of §6.2.5 below. But first, let us investigate what exactly 'taking tense seriously' means—that is, let us consider A-theory.

6.2.2 The Varieties of A-Theory

The distinction between A-theory and B-theory goes back to John McTaggart's famous essay on time [McTaggart 1908], in which he argued that time is 'unreal'.

[281] Merricks even mentions the 'tensed nature of predication' [Merricks 1994, p. 170]. For another defense of endurantism by way of 'taking tense seriously', see Zimmerman [1998].

Nowadays, the argument is often taken not to yield the 'unreality of time' but rather the 'unreality of tense', which is why it interests us at this point: it contains the thesis that A-theory is incoherent, which, given that A-theory is so closely related to the second-gear form of predication we wish to defend, would seem to imply that the second conceptual gear is *as such* incoherent. If so, we would, given our rejection of first-gear simulations of time, have to agree with McTaggart and conclude that not just tense but time as such indeed is 'unreal'. Luckily, however, McTaggart's argument is at best inconclusive—there are versions of A-theory that escape its threat. Unfortunately, some of these versions of A-theory, despite successfully rescuing tense from the McTaggartian threat of inconsistency, still rest upon a purely first-gear conceptual basis—in particular, on an eternalist understanding of time as a fourth dimension, as a time line. Instead of offering a directly first-gear understanding of time, like B-theory does, these versions of A-theory provide a first-gear conception of time *by* offering a first-gear understanding of tense. In that respect, such versions of A-theory are like the versions of endurantism we have discussed above: they amount to alternatives to the 'standard' first-gear approach to time in terms of B-theory and perdurance.

For reasons that will become apparent as we go along, our discussion of A-theory divides into two parts, which are separated by an investigation into the issue of eternalism versus presentism. In this first part of our discussion on time, we merely introduce the relevant sorts of A-theory and show how they rely on an eternalist background picture. Then follows what we have to say on eternalism and presentism in §6.2.3, after which we move on to argue, in §6.2.4 below, that reliance on eternalism enforces a first-gear understanding of the tenses: the tenses are, in effect, treated as modifiers operating on a basic stock of tenseless propositions—that is, the tenses are reduced to first-gear feature predication, hence not taken as a *sui generis* form of predication.

Now, let us start with McTaggart:

> Positions in time, as time appears to us *prima facie*, are distinguished in two ways. Each position is Earlier than some, and Later than some, of the other positions. And each position is either Past, Present, or Future. The distinctions of the former class are permanent, while those of the latter are not. If M is ever earlier than N, it is always earlier. But an event, which is now present, was future and will be past. . . .
> I shall speak of the series of positions running from the far past through the near past to the present, and then from the present to the near future and the far future, as the A series. The series of positions which runs from earlier to later I shall call the B series. [McTaggart 1908, p. 458]

McTaggart's argument runs, roughly, thus: time is real only if there is real change; there is real change only if the A-series exists; but the A-series is inconsistent—hence it does not exist; hence time is unreal. The thoughts underlying the argument are the following. Real change only exists if something changes in its properties. The A-properties are the ones that change, since they are not permanent while the B-relations are. Thus, change only exists if the A-series exists. But the A-series, as McTaggart thinks, entails that one and the same moment turns out to be both future and present and past, which is impossible because these properties are mutually exclusive. Hence the A-series is inconsistent, hence it does not exist. At best, then, the B-series exists, but that series is characterized only by permanent determinations, hence does not involve change, and therefore does not yield real time.

The argument has been extensively discussed throughout the last century, and opinions as to its cogency and value diverge widely.[282] Most people find that, whatever the cogency of McTaggart's reasoning, the way he set up the issue is useful. That is why, in the contemporary debate, there still is a division between A-theorists and B-theorists.

A-theorists 'take tense seriously', while B-theorists think of time in tenseless, 'permanent' terms. The latter think their account of time is adequate because time *is* the B-series, and in that respect they disagree with McTaggart (and with us). Moreover, they take themselves to be in a position to say everything an A-theorist wants to say as well: A-type statements are grounded in B-type statements, so the thought goes.[283]

The analogy with space is useful here. To make sense of your statement 'I am reading *here*', it is enough to point to the location where you are reading. No 'metaphysical here' is required to make sense of such spatially perspectival truths, and similarly, no 'metaphysical now' is required to make sense of your temporally perspectival statement 'I am reading *now*'. In both cases one can turn the seemingly perspectival truths into non-perspectival truths simply by including explicit reference to the relevant point in space or time from an 'absolute' or 'eternal' perspective. Quine, for instance, writes:

[282] See, e.g., Broad [1938], Prior [1967, ch. 1], Dummett [1960], Mellor [1981, 1998], Lowe [1992], Baldwin [1999], Dyke [2002], Zimmerman [2005], Fine [2005d], and Falvey [2010]. Broad, for instance, calls the argument at one instance a 'philosophical "howler" ' [p. 136], to which Zimmerman agrees, saying that this 'must be [McTaggart's] worst argument' [p. 401]. Dummett, on the other hand, says that 'his argument is not the trivial sophism which it at first appears' [p. 504]. We will mainly discuss Fine's similarly sympathetic reading of McTaggart.

[283] In Williamson's words (see the quote on p. 211 above), B-theorists thus reject *radical* transience and change.

> Any casual statement of inconsequential fact can be filled out into an eternal sentence by supplying names and dates and cancelling the tenses of verbs. Corresponding to 'It is raining' and 'You owe me ten dollars' we have the eternal sentence 'It rains in Boston, Mass., on July 15, 1968' and 'Bernard J. Ortcutt owes W.V. Quine ten dollars on July 15, 1968', where 'rains' and 'owes' are to be thought of now as tenseless. [Quine 1970, p. 13]

As we observed, B-theory comes down to the view that first-gear feature predication suffices for everything.[284] And indeed, the various B-theoretic views on persistence we discussed above (perdurantism and some versions of endurantism) amount to different ways of spelling out exactly what the form of such eternal, B-theoretic statements is—in particular, to different views on where the time reference should be placed. One could take them, thus, to provide different analyses of Quine's eternal sentences.

We should make two brief remarks on B-theory and tense. First, consider Quine's favored type of sentences. He calls them *eternal* sentences. If one takes 'eternal' to mean 'omnitemporal', the tenses are not being canceled but rather generalized, to yield, e.g., 'Ortcutt *owed, owes, and will owe* Quine ten dollars on July 15, 1968' (adopting Quine's second example). Such a sentence is true independently of the time and place at which it is uttered, but it is not what Quine has in mind: the tenses are still in there. We should understand the B-theorist's move not as replacing a *particular* temporal perspective (was, is, *or* will be) by a *general* such perspective (was, is, *and* will be), but rather as abandoning temporal perspective *entirely* in favor of a truly tenseless, i.e., *atemporal* form of predication.

Second, there seems to be consensus amongst contemporary 'new' B-theorists that they *do* 'take tense seriously', in a way. They agree with those objecting to replacing tensed truths by tenseless ones, but only insofar as the contents of our propositional attitudes are concerned: *these* are then taken to be irreducibly tensed, without this implying that a B-theoretic, tenseless description of reality is false or incomplete.[285] From our point of view, taking tense seriously in this non-committal

[284] There are other ways of resolving the perspectival element—e.g., by way of 'token-reflexivity': your statement 'I am reading now', for instance, is true just if your reading is simultaneous with the statement itself. This sort of reduction of A-properties to B-relations has been defended, e.g., by Russell [1919, 1941], D. C. Williams [1951, p. 463], Ayer [1956a, pp. 152–3], and Smart [1963, pp. 132–42]. Smart adopts the proposal from Reichenbach [1947, §§50–1].

[285] See, e.g., Lewis [1979a], Oaklander [1991], Oaklander and Q. Smith [1994], and Mellor [1998]. Zimmerman [2005] uses this kind of view for his argument that A-theory has to involve more than just 'taking tense seriously'. We may understand this change in the B-theorist's strategy to illustrate the broader movement from philosophy of language to metaphysics. Whereas earlier the issue between A-theory and B-theory concerned the question whether or not the tenses could be analyzed away *from our language*, now the issue is taken to concern whether or not the tenses play any role in the truth-makers for our (perhaps irreducibly tensed) true thoughts and utterances.

manner is typical for (Modern-picture) metaphysical realism, which separates reality as it is *for us* (in this case, tensed) from reality as it is *in itself* (in this case, tenseless). Most philosophers understand 'taking tense seriously' to mean more than just taking it seriously for the way we think—they intend to take tense *metaphysically* seriously. That is, they are interested in tense not merely as a linguistic phenomenon, nor as a feature of the way we think, but as a feature of reality itself (though for us this latter distinction is problematic—see §1.3).

Given our description so far, it is clear that B-theorists are committed to *eternalism*: holding that the temporality of temporal truths consists in their containing a reference to a temporal location, in some way or other, presupposes a view on reality as encompassing a complete spatiotemporal manifold. Such a manifold can be described in the atemporal way characteristic of B-theory. However, just like B-theory can be further spelled out in either perdurantist or endurantist terms, so eternalism admits not only of a B-theoretic but also of an A-theoretic rendering, as we shall soon find out.

A-theorists hold that there *is* a metaphysically special 'now', which, on their view, is precisely wherein time differs from space. They claim, for instance, that replacing the tensed *it is raining* with the eternal *it is*$_{tenseless}$ *raining at t*, is not an innocent rephrasing of essentially the same statement: what goes missing is that the mentioned time *t* is *now*—not merely in the relative sense of being simultaneous with our contemplation of the statement, but in an absolute sense.

There are different ways of spelling out this general idea. Arguably, the best-known versions of A-theory combine A-theory with presentism.[286] For present purposes, however, we prefer to first concentrate on A-theory as such, that is, apart from the presentism/eternalism opposition. Only by doing so can we investigate in which ways A-theory can be formulated on an eternalist basis, and how that is connected to first-gear metaphysics.

We follow Kit Fine's exploration of A-theory, since his rendering helps to bring to light the connection with eternalism and first-gear thinking. Like us, Fine chooses to focus exclusively on the issue of A-theory (what he calls 'tense-theoretic realism'), the issue of whether or not reality includes fundamentally tensed facts, quite apart from the issue of presentism versus eternalism. He writes:

> Presentism . . . is the view that only presently existing things are 'real' in some or another sense of the term. . . .

[286] See, for instance, Prior [1957, 1967, 1976], Geach [1967], Chisholm [1990a,b], McCall [1994], Lowe [1998, ch. 4], Merricks [1999], Hinchliff [2000], Crisp [2004], Markosian [2004, 2013], and Zimmerman [2005, 2008].

> Tense-theoretic realism, by contrast, is the view that reality is tensed; reality comprises tensed facts . . . . [Fine 2005d, p. 298–9]

Given this focus, Fine undertakes a reformulation of McTaggart's argument in terms of the following inconsistent tetrad:

> [*A-Theory*] Reality is constituted (at least, in part) by tensed facts.[287]
>
> *Neutrality* No time is privileged, the tensed facts that constitute reality are not oriented towards one time as opposed to another.
>
> *Absolutism* The constitution of reality is an absolute matter, i.e. not relative to a time or other form of temporal standpoint.
>
> *Coherence* Reality is not contradictory, it is not constituted by facts with incompatible content. [ibid, p. 299–300]

Fine's formulation of A-theory is quite straightforward. It involves talk of facts 'constituting' reality, which should be understood to mean that they obtain and are fundamental (roughly). *Neutrality* is more difficult. It is motivated by the thought that the present moment is only special as long as it *is* present, but no longer when it recedes into the past: there is no *one* privileged moment to which all the tensed facts are oriented. *Absolutism* is intended in a very strong sense: rejecting it means rejecting the idea of reality as being one, and results in a form of relativism (as we will see). And, finally, *Coherence* states at least that reality does not contain inconsistent facts—but it is rather difficult to understand what exactly the notion of coherence Fine has in mind amounts to, apart from there being no contradictory facts.[288]

These theses indeed form an inconsistent tetrad, and they do because of McTaggartian reasons: supposing that there is change, hence incompatible tensed facts obtaining at different times (by *A-Theory*), reality encompasses all of these incompatible facts impartially (by *Neutrality*). Moreover, they are then all part of the *one* reality (by *Absolutism*), and thus we come into conflict with *Coherence*: reality is incoherent. The argument against A-theory, now, rests upon holding on to *Neutrality*, *Absolutism* and *Coherence*, which forces one to reject *A-Theory*. The result is a rejection of tensed facts as basic building blocks of reality—in other words, B-theory.[289]

[287] We have made the terminological decision to relabel this thesis 'A-theory', instead of Fine's original label 'Realism', for reasons of clarity.

[288] What makes it difficult to understand Fine on this particular point is his insistence that a violation of *Coherence* does not amount to there being true contradictions (i.e., to robust dialetheism of the sort defended by Graham Priest, who claims that change and time involve true contradictions—see Priest [2006, chs. 11, 12 and 15]). Fine is very clear about this himself—see Fine [2005d, p. 282].

[289] See [ibid, §3] for a more elaborate development of the argument, and [ibid, §4] for a more sophisticated version that evades certain problems concerning the notion of 'constitution' involved. As an aside, notice that no contradiction ensues if reality is 'so boring that the same tensed facts hold at every single time' [ibid, p. 272]. (See fn. 308 on p. 247 for a critical note on this 'possibility'.)

Given this rendering of McTaggart's argument, it is clear that the A-theorist has three options. He may give up *Neutrality*, which is the usual way to go, yielding what we call *standard A-theory*.[290] This is the view we may associate with Arthur Prior, with whose views on and approach to time Fine sympathizes to a considerable extent.[291] Standard A-theory rejects the idea that all times are on a par, and thus singles out one particular moment—the present—which functions as the anchor point for all the tensed facts. Thus, supposing that you had breakfast yesterday, the tensed fact that you *had* breakfast is part of reality, because that fact is anchored in the present, whereas the associated tensed fact that you *are having* breakfast is not part of reality, because it is not anchored in the present (though it presumably was yesterday—and that is itself, again, a present-oriented fact). As we will see, there are radically different varieties of standard A-theory.

But the A-theorist may also hold on to *Neutrality*, and thus accept that both the fact that you *had* breakfast and the fact that you *are having* breakfast belong to reality, even though it is false that you are having breakfast *now* (assuming that you are not having breakfast whilst reading this chapter). Instead, he may reject *Absolutism*, yielding a position which Fine calls *external relativism*, but he may also give up *Coherence*, yielding a position which Fine calls *fragmentalism*. Let us briefly look into both of these options, which we view as different versions of *non-standard A-theory*.

The external relativist claims that there is not one reality but rather a multitude of realities, where each such reality is related to a certain temporal perspective. The tensed fact that you *are having* breakfast is part of one reality, albeit relative to yesterday's breakfast time, while the (incompatible) fact that you are *not* having breakfast is part of another reality, which is relative to the current moment (and which also includes the fact that you *had* breakfast). It is important that the relativity at play is not explained away in B-theoretic terms, or by reference to some underlying absolute notion of reality: reality is '*irreducibly* relative' [Fine 2005d, p. 279]. In a way, then, we may say that the external relativist agrees with the B-theorist that tensed truths are perspectival truths, but claims that the B-theorist is wrong in assuming that the perspective can be resolved. Instead of reducing time to an ordering on tenseless B-facts, the external relativist takes time to consist in a multitude of realities, and thereby manages to keep the facts constituting these various realities tensed.

[290] Fine calls it 'standard realism', but we find our label to provide a more tractable terminology in the context of our discussion of time.

[291] See, e.g., Prior [1957, 1967, 1968b, 1976], and of course Prior's posthumously published collection of essays entitled *Worlds, Times and Selves*, which Fine put together and wrote an extensive postscript for [Prior and Fine 1977]. We cast a closer (and critical) look at Prior's views below, in §6.2.4.

The fragmentalist, on the other hand, claims that there is one 'über-reality' which incorporates all the incompatible tensed facts—both that you are having breakfast and that you are not having breakfast, say. This über-reality decomposes into a multitude of fragments which are each internally coherent. Again, the incoherence involved is not to be explained away by reference to some underlying coherent notion of reality: reality is '*irredeemably* incoherent' [Fine 2005d, p. 281]. As in the case of the external relativist, the fragmentalist avoids reducing time to an ordering on atemporal B-facts, and instead takes time to consist in a multitude of fragments that make up the incoherent über-reality. He thereby manages to keep the facts that make up these various fragments tensed.

There is a close connection between the external relativist and the fragmentalist: the fragments of the latter correspond to the separate realities of the former, although it seems easier for the fragmentalist to allow for overlapping fragments, than it is for the relativist to endorse overlapping realities. Fine writes:

> [B]oth are concerned, in their own ways, to deny the existence of a single coherent reality. But the relativist denies that it is single, while the fragmentalist denies that it is coherent. [Fine 2005d, p. 281]

Fine presents a number of arguments to the effect that one of these non-standard A-theories is to be preferred to standard A-theory (with a slight preference for the fragmentalist option). One of these arguments is of particular interest to us: it concerns the *passage of time*. As John Norton, a philosopher of physics, writes:

> Time passes. Nothing fancy is meant by that. It is just the mundane fact known to us all that future events will become present and then drift off into the past. Today's eagerly anticipated lunch comes to be, satiates our hunger and then leaves a pleasant memory. ... Time really passes. It is not something we imagine. ... Our sense of passage is our largely passive experience of a fact about the way time truly is, objectively. The fact of passage obtains independently of us. Time would continue to pass for the smoldering ruins were we and all sentient beings in the universe suddenly to be snuffed out. [Norton 2010, p. 24]

It is the dynamicity of time—as opposed to space—which constitutes an important reason for many to endorse some form of A-theory, and this is true in particular for us: we wished to get away from the atemporal conception of persisting objects presented by first-gear views on persistence, and towards a view on which the very same banana can be first green and then yellow, as the result of a process of ripening. It is thus helpful for our purposes to consider how the various forms of A-theory Fine describes fare with respect to this feature of time.

Let us start with the standard A-theorist. Fine writes:

> [G]iven a complete tenseless description of reality, then what does [the standard A-theorist] need to add to the description to render it complete by his own lights? The answer is that he need add nothing beyond the fact that a given time t is present .... But then how could this solitary 'dynamic' fact ... be sufficient to account for the passage of time? ... [H]is conception of temporal reality, once it is seen for what it is, is as static or block-like as the [B-theorist]'s, the only difference lying in the fact that his block has a privileged centre. Even if presentness is allowed to shed its light upon the world, there is nothing in his metaphysics to prevent that light being 'frozen' on a particular moment of time. [Fine 2005d, pp. 286–7]

It becomes evident, at this point, that the version of standard A-theory Fine here considers comes down to what is often called *moving spotlight theory*. Here is Broad's original image of that view:

> in a fixed direction, ... the characteristic of presentness [is] moving, somewhat like the spot of light from a policeman's bull's-eye traversing the fronts of the houses in a street. What is illuminated is the present, what has been illuminated is the past, and what has not yet been illuminated is the future. [Broad 1923, p. 59]

Moving spotlight theory is a quite minimal way of endorsing standard A-theory: it consists in adopting the eternalist picture of reality that underlies B-theory, namely that of reality as containing all the tenseless facts about all the different moments, and adds the tensed fact that a certain moment is *the present moment*.[292] The addition of this 'moving spotlight' to the B-theoretic framework is supposed to account for the passage of time—but, as Fine rightly remarks, adding tensed facts only provides the spotlight theorist with a spotlight, not with its movement. What is remarkable about this form of A-theory, then, is that it *first* accepts the B-theoretic/eternalist construal of time as just another dimension to the spatiotemporal manifold that constitutes reality, and *then* seeks to re-add time as a dynamic feature of reality by adding a moving spotlight. The situation is similar to that in the case of the endurantisms we surveyed in §6.2.1 above: there, we noticed that they *first* adopt a B-theoretic/eternalist understanding of temporal truth, and *then* seek to re-add the diachronic identity of persisting things by building the temporal reference of such B-theoretic truths into the predicates rather than into the objects.

The question Fine poses is, thus, indeed an important one for the moving spotlight theorist: he may *say* that the tensed fact he adds to the tenseless description is 'dynamic', i.e., that the spotlight 'moves', but that is merely to restate what is required,

[292] Defenders of something like the moving spotlight theory include Russell [1915], Q. Smith [1993, 2002], and Craig [2000]. Moreover, Skow [2009] contains an attempt to make moving spotlight theory compatible with special relativity—see also §6.2.6 below.

not to account for it. The tensedness of the fact itself does not seem to do the trick—which, we should add, indicates that there may be something fishy about the kind of tensedness involved.

Moreover, even if the moving spotlight theorist somehow manages to get his spotlight moving, a problem remains. For, as Broad observed long ago:

> [T]he lighting of the characteristic of presentness now on one event and now on another is itself an event, and ought therefore to be itself a part of the series of events, and not something that happens to the latter from outside. [Broad 1923, p. 60]

In other words, the facts about the motion of the spotlight are where the real temporal action is, so to speak, yet this temporal action is precisely *not* included in the time line. The moving spotlight picture thus founders on the, perhaps unintentional, introduction of a second temporal dimension. Hence the critical question, also raised by Broad, what the speed is with which the spotlight moves over the time line: how many time line–time-units does the spotlight traverse in one spotlight–time-unit? Does the spotlight ever accelerate or decelerate? Does it ever pause for a while on one spot of the time line, or even move backwards? The time line over which the spotlight moves turns into a space-like dimension for the spotlight to move over, and the movement itself takes place in this silently introduced extra temporal dimension.[293]

With respect to standard A-theory, we conclude, first, that simply *adding* tensed facts to an otherwise tenseless reality does not suffice for an account of time, and, secondly, that having the anchor for those tensed facts—the present moment—'move' over the time line results in a confused doubling of time. The moving spotlight version of standard A-theory fails as an account of time, and the reason why it fails is its adherence to the time line, that is, to an eternalist picture of time. It is, of course, possible to drop the underlying eternalism and endorse a presentist version of standard A-theory. We postpone discussion of such a view until we have had occasion to consider presentism in §6.2.3 below.

So, let us now turn to the non-standard forms of A-theory. How do they fare with respect to time's passage? Fine writes:

> For the external relativist, each time is objectively present at that time: at each time t, reality is constituted by the absolute fact that t is present .... And for the frag-

[293] This criticism can also be found in D. C. Williams [1951, pp. 463–4]. See Markosian [1993] for a recent discussion of the argument. Skow [2009] argues that this extra temporal dimension can be cashed out in terms of tensed truths concerning the 'real' time line—but see Pooley [2013, §IV] for convincing criticism, leading to the conclusion that the moving spotlight view endorses an eternalist take on ordinary time and a presentist take on the extra temporal dimension that underlies the movement of the spotlight.

> mentalist, each time t is objectively present *simpliciter*—i.e. reality is constituted by the absolute fact that t is present .... But in either case, presentness, in so far as it is a genuine feature of reality, applies equally to all times. Presentness is not frozen on a particular moment of time and the light it sheds is spread equitably throughout all time. [Fine 2005d, pp. 287–8]

This is the result of retaining *Neutrality*: for both the relativist and the fragmentalist, all times are on a par, none is singled out as special. Hence presentness applies to all of them equally.

Now, Fine of course knows that such an equal distribution of presentness over all times is still quite far away from an adequate account of the passage of time, but he nevertheless takes it to be an improvement in comparison to the situation of the standard A-theorist:

> [A]t least, on the current view, there is no obvious impediment to accounting for the passage of time in terms of a successive now. We have assembled all the relevant nows, so to speak, even if there remains some question as to why the relationship between them should be taken to constitute a genuine form of succession. [ibid, p. 288]

In terms of the moving spotlight, we could interpret Fine here as pointing towards the following consideration. For the moving spotlight version of standard A-theory, the spotlight is frozen on one particular moment of the time line, given that there is no time for the spotlight to move in outside of the time line itself. The non-standard A-theories, on the other hand, in effect partially *incorporate* the movement of the spotlight into their conceptions of the time line, by including the fact that the spotlight shines on a particular moment into that moment itself (conceived of either as one of the many realities or as one of the many fragments).

It is striking, however, that Fine does not notice the weakness in these considerations. We saw that the moving spotlight theorist starts out by accepting the static, eternalist picture of temporal reality, and then attempts to re-add dynamicity by turning on his moving spotlight. By incorporating the presentness of each moment into the moment itself, the non-standard A-theorists simply give up on the dynamicity that motivated the spotlight view, and hence the addition of tensed facts, in the first place. Indeed, adding presentness to each moment sounds very much like B-theory, where each moment is, after all, present relative to itself. True, the non-standard A-theorists reject the tenseless, B-theoretic understanding of temporal truth, and therefore the presentness of each moment is, on their view, not relative but absolute—but they still accept the underlying, static eternalist picture, and even to a

further extent than the moving spotlight theorist.[294]

It is clear, then, that the moving spotlight theorist and the non-standard A-theorists are all roughly in the same boat when it comes to the metaphysics of time. Simply adding tensed facts is not enough to re-add dynamicity to the underlying eternalist picture. Adding many nows is just as hopeless as adding one now—the spotlight remains frozen, no matter how wide or narrow its beam. We can say, thus, that the relativist and the fragmentalist are in effect subscribing to an alternative version of moving spotlight theory. That alternative version involves an (unsuccessful) attempt at unifying the spotlight theorist's two time dimensions, that is, an attempt at incorporating the movement of the spotlight itself into the time line.

What makes all the versions of A-theory Fine discusses, while considering the passage of time, unsuitable as accounts of time is that they do not question the eternalist picture on which B-theory rests. They assume that reality is temporally encompassing, yet they wish to hold on to the reality of tense, by way of insisting that the time line is not populated by tenseless, B-theoretic facts but rather by tensed, A-theoretic facts. This is clear especially in the case of the non-standard A-theories: the starting point is the eternalist collection of all temporal facts, but instead of construing them as tenseless B-facts, that peacefully coexist together in the same big container called 'reality', they insist on construing them as tensed A-facts, which enforces either a relativization or a fragmentation of reality.

We conclude that A-theory as such is not enough to distance ourselves from a first-gear understanding of temporal reality: it can be combined with an eternalist picture that is shared with B-theory. What we need to do is, therefore, to question that eternalist picture. Once we have done so, we will be in a position to show that, despite appearances, the way Fine's versions of A-theory conceive of temporal truths does *not* involve second-gear predication, which is truly temporal, but rather rests on a foundation of first-gear feature predication which is supplemented by surrogate 'tenses' as an external addition. Second-gear temporal predication is incompatible with eternalism, hence the understanding of the tenses involved in eternalist A-theories cannot be the one that is at home in the second conceptual gear.

[294] To prevent confusion: even the external relativist takes presentness to be a feature that belongs to each moment in an absolute manner. The relativity of the external relativist does not consist in some fact's being *present* relative to some moment but not relative to some other moment. Rather, the relativity in question concerns whether or not the facts belong to reality *at all*. Yesterday's presentness belongs to yesterday's reality, and is simply not to be found in today's reality, which contains today's presentness.

6.2.3 From Negative to Positive Presentism

Throughout our discussions both of theories of persistence and of A-theory, we repeatedly observed that the picture of a time line which can be looked at from an atemporal or 'eternal' perspective is the cause of much trouble. We concluded that a viable version of endurantism, which does not collapse into a mere variant of a first-gear simulation of persistence, requires that we stop rendering temporal truths atemporally, that is, that we abandon B-theory for good. Endurantism, insofar as it interests us here, thus requires A-theory: just stating that things are 'wholly present whenever they exist' is not enough, for the 'whenever they exist'-part may well be understood in B-theoretic, atemporal terms. We subsequently concluded, after surveying some variants of A-theory, that taking tense seriously does not get us any further: replacing the tenseless B-facts that occupy the eternalist time line with tensed A-facts merely complicates the eternalist picture. Tense does not yield anything like the dynamicity we were hoping to gain, the passage of time remains a myth for these A-theorists as it was for B-theorists[295]—unless we manage to divorce A-theory from the eternalist picture that it seems to be compatible with. Just stating that 'reality comprises tensed facts' is not enough, for the 'tensed facts' may well be understood such as to be compatible with their inhabiting an eternalist, static reality, as Fine's versions of A-theory illustrate.

Now, even presentism itself, being usually understood in opposition to eternalism (or even defined as its opposite), can, perhaps surprisingly, be endorsed against the background of an eternalistic picture. Such a form of presentism results if the opposition of eternalism to presentism is presented as follows. Take the picture of the time line, and mark the dot which is the present moment. Now ask: shall we take reality to include the *entire* time line? Or shall we say that reality includes only that one dot which we marked as 'the present'? Eternalism, on this presentation, amounts to the first option, while presentism amounts to the second option. We call presentism, thus defined, *negative presentism*.

Negative presentism is unable to account for time at all: if we erase the rest of the time line, it seems that we are stuck with a reality that is purely spatial, and not temporal. Negative presentism thus does not get us any further in our search for a second-gear conception of time: it is built not on a rejection of the eternalist picture but rather accepts it and *then* modifies it. Thus, as we should expect, the characteristic

[295] D.C. Williams's famous essay entitled 'The Myth of Passage' [D. C. Williams 1951] treats the idea of passage as an attempt to resist what he calls the 'dimensional view of time' [e.g., pp. 465 and 469], that is, eternalism.

problems we ran into in our discussion of eternalism-based versions of A-theory still apply to negative presentism. Most importantly, time's passage remains a mystery, for nothing in the negative presentist's metaphysics ensures that his reality, confined as it is to just one moment, changes. Reality collapses into one static snap-shot and is therefore entirely atemporal. This is the counterpart, for negative presentism, of the 'frozenness' of the spotlight we observed in the case of Fine's eternalist version of standard A-theory (moving spotlight theory).

Like moving spotlight theory, negative presentism is usually accompanied by the (mere) claim that time passes, that things change. In that combination, it seems to be the standard version of presentism—at least, according to those opposing it. Fine, for instance, contrasts the issue of A-theory he is concerned with with presentism (what he calls 'ontic' presentism) as follows:

> Ontic presentism is an ontological position; it is a view about what there is. [A-theory], on the other hand, is a metaphysical rather than an ontological position; it is a view about how things are, quite apart from what there is. ... Moreover, ontic presentism is a negative view; it excludes certain things from what there is. [A-theory], on the other hand, is a positive view; it includes certain ways of being in how things are. ... It is readily possible for [an A-theorist] not to be an ontic presentist. Indeed, he may endorse a full ontology of things past, present, and future .... He merely insists that some of the facts (if not all) should concern how things presently are. I am inclined to think that this version of [A-theory] is much more plausible than the usual version, in which only present things are taken to exist; and it is a shame that a one-sided conception of the presentist issue has prevented philosophers from taking it more seriously. [Fine 2005d, pp. 299-300]

Fine rightly takes *negative* presentism to stand in the way of a proper understanding of time. He fails to notice, however, that what is generating the problem is the underlying eternalist picture. He thinks that, instead, the problem lies in the preoccupation with the 'ontological' question as to the existence or non-existence of past and future things. That is why he sees no problem in adopting the eternalist picture even in his own investigations into A-theory, as we have seen.[296]

Negative presentism helps us to pin down what it is about the eternalist picture of a time line that causes trouble. It makes clear that the root of the problem is not the all-inclusiveness of the time line. Rather, the problem lies in the way in

[296] We disagree with Fine on a different but related point as well: the distinction between metaphysics and ontology he draws. On our view, the questions of *how* things are and *what* things are, are far from being as independent as Fine here suggests. That should be clear from our introduction of the various conceptual gears in chapter 4: if *what* there is includes only first-gear entities, then *how* they are will be determined by purely first-gear forms of thought, while if *what* there is, includes second-gear physical objects, then *how* they are will be determined by second-gear forms of thought. And it is only the former that lack temporality.

which the eternalist picture ties truths or facts (be they tensed or tenseless) to times: every moment comes with its own set of truths. The moments are, thus, conceived of rather like possible worlds: alternative ways the world might be, ordered in a 'temporal' series. Or, to be more precise, alternative ways an *instantaneous*, that is, *atemporal* world might be. Time is conceived of as a series of entirely separated, atemporal instants. A mere collection of global world-states. The Humeanism in this conception is obvious: all the instants are 'loose and separate'.

That is why the eternalist picture naturally invites a B-theoretic conception of temporal truths: to describe what reality is like at one particular instant, it is enough to use atemporal, first-gear feature predication: time comes in only as one considers that instant in its relations to other instants on the time line. It is not surprising, therefore, that insisting on tensed facts against such an eternalist background picture does not make any difference when it comes to the metaphysics of time—just as we have seen above.

Many of the well-known charges against presentism in fact aim at negative presentism. Take, for instance, the truthmaker argument: there were dinosaurs, so the thought goes, but there is nothing in the present moment which makes true that truth about the past.[297] Or, as another version of the argument goes, take your admiration for Aristotle's wit: neither Aristotle nor his wit are presently existing, so how can you stand in the relation of admiring to either of them? Does not a relation presuppose the existence of its relata?[298]

We submit that any version of presentism which is vulnerable to arguments such as these is wrong-headed from the start. The eternalist times-as-locations picture is still in place; it is just that reality is being restricted to only one of these locations, which makes all the other locations inaccessible. And that is what underlies arguments of this kind. *What the truthmaker arguments show, then, is that there is no difference between this presentist, A-theoretic picture and an eternalist, B-theoretic picture of a temporally thin reality.*

Let us elaborate on this point. The negative presentist does not give up the eternalist's conception of the way in which temporal truths are allocated to their proper moments. All he gives up is the extendedness of the time line. His time line

[297] If you think fossils can do the job, then replace the example with one involving a past contingency that has left no traces at all in the present moment.

[298] The truthmaker argument, and similar arguments, are discussed by, e.g., Prior [1968a], De Clerq [2006], Bourne [2006], Fiocco [2007], Caplan and Sanson [2011], Baron [2012], Torrengo [2013], and Asay and Baron [2014]. Also, see Dummett [2004] and Markosian [1995, 2013] for explorations of what one might read as negative presentist views which allow for the possibility that the past is open just as the future is, since past truths, like future truths, can only have present truthmakers.

shrinks to the size of one single moment: the present. He alleges this present moment to be characterized by *present*-tensed truths. An eternalist, B-theoretic but temporally thin reality would incorporate atemporal truths concerning the only 'moment' that exists. These truths are tenseless, of course. Now, it is highly questionable whether such a reality incorporates time at all—after all, we could then declare the two-dimensional Euclidean plane to be just such a 'temporal' reality. Hence, it is equally questionable whether the negative presentist, A-theoretic picture we are considering incorporates time at all, hence also whether the alleged present-tensed truths that characterizes the 'present' are really present-tensed and not simply atemporal. The point becomes especially vivid in the context of Bourne's ersatzist version of negative presentism. He writes:

> I propose we construct times using maximally consistent sets of *u*-propositions, which intuitively we can see as those *u*-propositions that are true at that time. These propositions I take to give a complete, maximally specific, description of what is true at that time. ... [These times] need to be ordered by an 'earlier than' relation ..., in order for the ersatz time series to be structurally similar to a real time series, so it can be taken to be a sufficient substitute. [Bourne 2006, pp. 53–4]

The '*u*-propositions' are here taken to be the present-tensed propositions. In effect, then, Bourne proposes to replace the eternalist's time line with an ersatzist time line consisting of a series of maximally consistent sets of present-tensed propositions. One point on this substitute time line is occupied by reality: the present moment. The earlier-later relation that holds between all the times (the ersatz times as well as the one real time) grounds the various past and future truths at the various moments. And, since the sets of propositions are abstract objects, they exist timelessly and hence can function as truthmakers for past (and, if you like, future) truths at the present moment.

This may perhaps solve the truthmakers problem. But obviously, it does not get us any further with regard to the metaphysics of time. While the eternalist has no need for the passage of time, since on his view the time line as a static dimension to reality is all we need, the presentist *does* need the passage of time, for otherwise his present moment is condemned to be 'eternally present'. And that is just the kind of picture Bourne's position threatens to collapse into.

For a change, then, we may agree with Lewis, who writes with regard to persistence on a view like Bourne's:

> The only intrinsic properties of a thing are those it has at the present moment. Other times are like false stories; they are abstract representations .... When something has different intrinsic properties according to one of these ersatz other

> times, that does not mean that it, or any part of it, or anything else, just has them—no more so than when a man is crooked according to the Times, or honest according to the News. This is a solution that . . . rejects persistence altogether. [Lewis 1986a, p. 205]

Lewis intends his line of thought to apply to presentism in general, yet it is clear that the only form of presentism that he is capable of contemplating *in first-gear terms* is negative presentism. To that extent we agree: negative presentism amounts to a rejection of persistence—and not only of persistence but of time altogether.

It is not surprising, then, that Bourne's position can easily be understood in entirely atemporal terms. If we take his *u*-propositions, which are supposed to be present-tensed, to be tenseless instead, we get the kind of view that Lewis has in mind: it is as if we add to a certain distribution of properties over the two-dimensional Euclidean plane a series of alternative 'ersatz planes', each with its own distribution of properties, ordered by an 'earlier-later' relation, in terms of which we can then talk about the 'past' and 'future' of the plane.

Presentists such as Bourne mistakenly think that presentism faces the mentioned problems involving truth-makers—and hence are mistaken when they think they can solve them, too (recall Merricks's claims to similar effect concerning endurantism and the problem of temporary intrinsics—see the quotes we provided in §6.2.1, p. 212). A proper understanding of presentism will not be like negative presentism, and will show that there are no problems of this kind.

The kind of presentism we seek is not equivalent with an eternalist, B-theoretic picture of a temporally thin reality. It will reject a basic stock of 'present-tensed' truths which can equally be read as tenseless truths. And it does so by rejecting the very idea of a distribution of truths over separated instants, as the eternalist picture has it.

The kind of presentism we propose, *positive presentism*, is based on its own, positive construal of temporal reality and of temporal truth, instead of being based on an eternalist picture which it then partly negates. Time is not a dimension to reality, it is not a collection of instants. Rather, it is a *sui generis* mode of being. Temporal existence—persistence—does not rest on a basis of atemporal, first-gear existence *at an instant*, and neither does temporal truth rest on a basis of atemporal, first-gear truth *at an instant*—nor on a basis of tensed truth *at an instant*.

The difference between negative and positive presentism consists in the former's being defined in terms of the time line, while the latter seeks to ground itself in a truly first-gear version of A-theory (like the version of endurantism we will sketch in §6.2.5 below). Such a version of presentist A-theory differs from the eternalism-

based, moving spotlight version of standard A-theory we discussed earlier in that it does not presuppose eternalism, and hence does not face the 'double times' problem that the spotlight theorist faces (because no time line is presupposed). In a way, then, the proposal is to keep the moving spotlight but drop the time line.

Still, one might think, Fine's worry as to the 'frozenness' of the present may be raised. What do we say? Interestingly, Oliver Pooley recently defended presentist A-theory against Fine's worry with the help of Prior's account of presentist passage. Change consists, on Prior's tense-logical understanding (on which more below), in the truth of tensed conjunctions like '*b* was *G* and *b* is not *G*'. Observe that 'Prior's birth was 54 years ago, and his birth is now not 54 years ago' has the very same form. It is a kind of change, then—and Prior writes, with regard to such a peculiar change:

> This last change, of course, is a case of precisely that recession of events into the past that we are really talking about when we say that time flows or passes .... [Prior 1968a, p. 9]

Thus, Pooley says, if we adopt Prior's tense-logical understanding of change, we inherit a perfectly fine conception of the passage of time. He writes:

> The present, tensed facts include, for example, facts to the effect that certain tensed propositions are not now true *but that they were or will be true*. One simply cannot accept all the present, tensed truths without accepting that what is true undergoes genuine change. [Pooley 2013, p. 330]

Below, we will develop an understanding of tensed facts that will do what Pooley takes Prior's conception of them to do: to provide an understanding of both change and the passage of time. That is, our positive version of presentism rests on such an understanding of tense—on a proper A-theory.

For now, the point of bringing in Pooley's use of Prior's understanding of passage is the following. Consider the way Pooley extends this understanding of passage to Fine's non-standard A-theories:

> Do non-standard views vindicate the passage of time? The first, obvious point to make is that everything the presentist said was true absolutely remains true relative to a particular temporal perspective. And everything the presentist maintained was always true remains true relative to every temporal perspective. Since time passes for the presentist, the same holds true, *as of any time*, of the non-standard view. One of the view's many perspectives is supposed to be our perspective so we can truly say (now) that time passes. [ibid, p. 335; emphasis added]

The non-standard A-theories incorporate instants in the sense we found problematic: the eternalist's time line is being chopped up to yield a multitude of relative realities

(or fragments), but it is there nonetheless. Insofar as Prior's tense-based understanding of change and the passage of time is successful in the case of presentist standard A-theory, it yields passage of time *for the presentist's reality*. Now, the one single reality of the presentist standard A-theorist *is identical to* one of the many realities (or fragments) of the non-standard A-theorist: exactly the same tensed truths obtain in both realities. So, contrary to what Pooley thinks, the respective generalization of Prior's account yields as many 'moving nows' as there are relative realities (or fragments).[299] By marrying the Priorian idea of understanding time in terms of the tensed form of predication with the idea of a multitude of temporal locations (i.e., realities or fragments), Pooley ends up with a host of different times: one time for each moment. It is clear that *that* is not what a non-standard A-theorist has in mind: he rejects such a Priorian understanding of time in favor of a (non-standard) eternalist understanding.

Thus, we may say that insofar as our presentist standard A-theorist conceives of the relation between tensed truths and the present moment in the way the non-standard A-theorists conceive of the relation between tensed truths and all their diverse moments (realities, fragments), *such* a presentist standard A-theory cannot be what we are looking for: it still rests on a positioning of temporal truths on a time line. In other words, one should take great care in explicating one's understanding of tense.

This illustrates the problem with understanding reality to include tensed truths *at an instant*. Instants are, at best, constructions out of tensed truths, not some location for them to reside in—as Prior writes:

> I find myself quite unable to take 'instants' seriously as individual entities; I cannot *understand* 'instants', and the earlier-later relation that is supposed to hold between them, except as logical constructions out of tensed facts. Tense logic is for me, if I may use the phrase, *metaphysically fundamental*, and not just an artificially torn-off fragment of the first-order theory of the earlier-later relation. [Prior and Fine 1977, p. 37]

Our proposal, then, is to conceive of temporal truths as full-blown second-gear truths, not merely as tensed truths in the way Fine, Pooley, and others propose—tensed truths, that is, which may be located at instants or moments. Presentism, on our view, is the thought that reality is inhabited by things to which second-gear forms of predication apply, that is, by persisting objects capable of undergoing change.

[299] The point becomes especially vivid when considering the open future, for then what is taken to be genuinely open relative to the present reality (say, whether or not there will be a sea battle tomorrow) is no longer open relative to some future reality—which seriously compromises the sense in which it is presently 'open'. See §6.2.6 below.

Reality indeed lacks temporal extension, not because it only has a minimal temporal extension, but because the very idea of temporal extension is mistaken. Reality is *in itself* temporal by being the container of second-gear objects and the tensed facts about them.

So, how does reality move from one moment to the next, then? What is time's passage? It is important that we be clear on what exactly this question demands: only if we tacitly presuppose the picture of the time line is there a serious question as to how reality 'moves from one moment to the next', how one collection of (tensed or tenseless) facts is replaced by a fresh collection. Instead, we adopt the Priorian take on passage as linked to change: time's passage consists in the persisting objects' being engaged in processes of various sorts. That is, we need to take into account the conception of causation as genuine production we discussed earlier (§6.1): the objects that exist are constantly active; they *produce* the next state of reality, thereby making true (in a very literal sense) new truths, and increasing the temporal distance to old truths. As said, what exactly this amounts to will be our theme below.

Change, then, plays a crucial role in our understanding of time. Causality and time are two sides of the same (second-gear) coin, so to speak—as is witnessed by the fact that aspect and tense both occur on the second conceptual gear but not on the first conceptual gear. Following Wiggins, we may thus say:

> Change is more primitive than any ontology of times. Let us not visit upon 'I shall soon be bald' the problems that come to be perceived in 'For some t, I am [timeless] bald at t and t is not too long after now.' [Wiggins 2001, p. 179; the bracketed insertion is Wiggins's]

With Prior and Wiggins, we propose, then, to understand the eternalist picture of a time line as *derivative*, not fundamental. We live in a reality that is temporal—otherwise it would not be possible to 'live' at all, we may even say—and that reality, by constantly changing, can be thought of as occupying a position on a time line, on which it is moving forward. The picture is to be understood in terms of time, and not *vice versa*. The time line is a spatial metaphor for time, not time itself. Of course, if we keep this in mind, the time line can be a very useful tool in the formal/mathematical study of many physical phenomena, as has been illustrated time and again by successful physicists.

One could formulate our positive version of presentism in the following way. Every attempt at 'accounting for' or 'analyzing' concepts like time, change, persistence, tense, and causation is mistaken. There is no 'accounting for' time, change, persistence, tense, and causation, since these concepts are the fundamental concepts

of a *sui generis* form of being, which we have sketched as the second conceptual gear in §4.3. At best, one can attempt to elucidate or explicate that conceptual gear. The basic second-gear concepts are, as Sebastian Rödl says, 'categories of the temporal' [Rödl 2012/2005]. Presentism, A-theory and endurantism are included in this conceptual gear *not* as separable ingredients that can be opposed to eternalism, B-theory and perdurantism, but rather as inseparable ingredients without which the entire second conceptual gear breaks down. The only real opposition there is, then, is the opposition between the second conceptual gear and various simulations thereof in first-gear terms.

This might seem to be a very unsatisfactory conclusion, and for those expecting an account of time in first-gear terms it inevitably is. For those seeking to make the shift from first- to second-gear thinking, however, it is the only way to go. We have seen, after all, that on all the major dimensions of the contemporary debate on time we discussed (presentism versus eternalism, A-theory versus B-theory, and endurantism versus perdurantism) first-gear thinking may be upheld, no matter how far away we move from its paradigmatic, eternalist/B-theoretic/perdurantist simulation of time. Again, the only opposition that makes a truly substantive metaphysical difference is the basic opposition between full-blown second-gear thought and first-gear simulation thereof.

Put in slightly paradoxical terms, then, the only viable account of time consists in giving up the search for an account of time. Yet this does not mean that nothing is to be said about time. For instance, we have already claimed that our understanding of time as a basic category of second-gear thought is a form of presentism: positive presentism. It is a peculiar form of presentism in that there is no possible form of eternalism with which it can be contrasted—at least, not as long as we stay within the second conceptual gear.

Given the fact that the versions of presentism, A-theory, and endurantism we wish to defend are to be thought of as interrelated aspects of the second conceptual gear, it is not surprising that our version of presentism, positive presentism, presupposes a proper A-theory—as we observed repeatedly. Let us thus proceed by showing in what way positive presentism is also a form of A-theory—and, again, a peculiar form of A-theory in that there is no possible form of B-theory with which it can be contrasted on the second conceptual gear. Surprisingly, what we need in order to arrive at this version of A-theory is to reject a very influential assumption made by Prior, one of the leading figures both of contemporary presentism and of contemporary A-theory.

6.2.4 Tense Logic Rests Upon a Mistake

We have seen that many A-theories *agree* with B-theory in that they include the picture of time as a time line, as a series of separate moments. First, we discussed A-theories that adopt eternalism: all the tensed facts, from all the different moments, are contained in reality; it is just that they are not construed atemporally as the B-theorist does. Instead, the A-theorist keeps them tensed. Thereby, the tensed facts float free, so to speak, from their proper temporal locations. The standard A-theorist chooses one moment, the present, and includes only those tensed facts that are oriented towards that moment. The passage of time is then understood as a shift of focus over the time line (the moving spotlight). The non-standard A-theorists do not restrict the scope of reality in that way; instead, they sort all the free-floating tensed facts into different realities (for the relativist) or into different fragments (for the fragmentalist). Then, we discussed a presentist version of A-theory we called negative presentism, and observed that taking reality to encompass only the present moment is still compatible with the underlying eternalist conception of time. And finally, we considered a presentist version of A-theory that rejects the eternalist picture and looks to a proper understanding of tensed truth for the accompanying understanding of the passage of time. Prior has aimed to develop just such an understanding of tensed truth, as we observed following Pooley (see p. 229 above).

Just making sure that the basic facts are tensed is not enough to make sure that one stays within the second conceptual gear. For if the understanding of tense involved allows for non-standard A-theories, as Fine and Pooley think, such tensed facts are still treated as being capable of occupying a location on a metaphysically fundamental time line. Accordingly, as we saw in the case of the negative presentist, it is surprisingly easy to read the present-tensed facts that characterize the present moment as tenseless facts instead.

We argue, here, that the understanding of tense on which all of these varieties of A-theory are based rests, in the end, on first-gear feature predication. As said, we are working towards an understanding of tensed truth that is sympathetic to what Prior wished his understanding to deliver—but we will have to amend his approach in decisive ways. To that effect, we first need to bring out an oddity in the way Fine (and Pooley), following Prior, conceives of tensed facts.

Consider, for instance, the fact that you were having breakfast yesterday morning. That fact obtains now. From the standpoint of yesterday morning, however, the fact that you *are* having breakfast obtained. The negative presentist and other standard A-theorists will have only the first of those two in their reality (since the present

moment is today); the non-standard A-theorists will have them both and put them into different realities or fragments. The oddity in this understanding of tensed truths, apart from their being treated as capable of occupying a temporal location, is that the two are taken to be two *different* facts.

This way of understanding tensed facts derives directly from Prior himself:

> In the logic of tenses, the ordinary statement-variables p, q, r, etc., . . . are used to stand for 'statements' in the sense in which the truth-value of a statement may be different at different times. [Prior 1957, p. 8]

Such 'statements' are, thus, tensed statements, *but restricted to the present tense*. The logic Prior constructed on the basis of such statement-variables makes use, in addition to the standard logical connectives, of the famous tense operators **F** (it will be the case that . . .) and **P** (it has been the case that . . .) and their duals. One of Prior's leading thoughts is that we can, obviously, only make statements and have thoughts that are true relative to the present moment, but that the tense operators make available to us truths relative to other times—if p is true relative to the present moment, then $\mathbf{P}p$ is true relative to tomorrow. However, the statement that is true now, viz., that *you are reading* (p), is a different one from the statement that is true tomorrow, viz., that *it was the case that you are reading* ($\mathbf{P}p$): p is simply not the same as $\mathbf{P}p$. That the first statement is part of the second makes them related in a certain way, but not identical.[300] Moreover, if we forget about the time line and restrict ourselves to what is true *now*, as a Priorian presentist should, we cannot even make statements about tomorrow's truths. If p is true now, the best we can do is state that tomorrow $\mathbf{P}p$ will be true—but that is just to say that $\mathbf{FP}p$ is *now* true.

We see, now, that Fine takes this non-identity very seriously: since the formula 'p' is not the same as the formula '$\mathbf{P}p$', the fact that p is not the same as the fact that $\mathbf{P}p$ either. Yet it is clear that what Prior had in mind is that '$\mathbf{P}p$' tomorrow will state *the same thing* that 'p' states now. It is just that he was not willing to yield to the B-theorists: they say that that 'same thing' can only be a *tenseless* fact of the form p-at-t (roughly). That is, Prior wants to resist interpretation of his tense-logical framework in terms of a B-theoretic, eternalist framework. To repeat part of the quote we presented before:

> I cannot *understand* 'instants', and the earlier-later relation that is supposed to hold between them, except as logical constructions out of tensed facts. Tense logic is for me, if I may use the phrase, *metaphysically fundamental* . . . . [Prior and Fine 1977, p. 37]

[300] Compare *you are reading* (p) with *it is not the case that you are reading* ($\neg p$), or with *you are reading or* $1 + 1 = 2$ ($p \vee q$).

We agree wholeheartedly with the spirit of what Prior writes: the time line, with its instants, is a construction (or, as we said, a spatial analogy), not time itself. Whether or not this means that 'tense logic' is 'metaphysically fundamental' is, however, precisely what we now wish to contest—though we undoubtedly agree at bottom with Prior on what he takes to be metaphysically fundamental about time (namely, that it is a basic category of the second conceptual gear), we have reasons to doubt that his tense logic captures that fundamental aspect of time—doubts raised by the use Fine and Pooley make of the way tense logic treats the tenses.

As we have seen, Fine is faithful to Prior's project: his take on A-theory is to view reality as constituted by tensed facts of just the forms to which Prior's tense logic gives rise—he thus takes tense logic to be metaphysically fundamental. But he forgets about the thought that our current statement that *you are reading*, p, says the *same thing* as does the statement we may make tomorrow by saying that *you were reading yesterday*, $\mathbf{P}p$. For he simply includes both p and $\mathbf{P}p$ in reality as separate facts. It is important to realize that *one can only construct non-standard A-theories by conceiving of tensed facts in this manner*.

A-theories that are based on this understanding of tensed facts in effect isolate the different moments from one another entirely. For consider: there is nothing in this conception of tensed facts that secures that the facts of different moments will cohere in the way one expects. Perhaps tomorrow's stock of tensed facts includes $\neg\mathbf{P}p$ even though today's stock includes p, and $\mathbf{FP}p$, and any other fact one might use *today* to state that tomorrow $\mathbf{P}p$ will hold. The only thing that tense logic can guarantee is that *at one and the same moment* the facts cohere as they should: that much is settled by axioms like $p \rightarrow \mathbf{FP}p$. The present moment, then, has its very own set of tensed facts, and these are all completely independent from the facts that obtained just a moment ago and that will obtain in just a moment. The non-standard A-theories make this temporal isolation very explicit; in the case of the standard A-theorist and the negative presentist it is less visible since they only accept the tensed facts that relate to the present moment. The point is, now, that this understanding of tensed facts in a way divorces the tensedness of the facts from time itself—and allows them to be placed on a time line in the way Fine and others do.

It may seem that we are being unfair to Prior here—haven't we tacitly slipped into talking in terms of instants housing sets of tensed truths, which is just the notorious eternalist picture Prior, like us, wants to get rid of? That is right—but it also shows that, insofar as Prior's take on tensed facts lends itself to such use, he has simply failed to reach his goal. We need to keep in mind, here, that Prior of course wants to

say that time passes, that things change—and he indeed provides an explanation of how that dynamical aspect of time ties in with tense logic. He writes:

> In tense logic the totalities of tensed propositions which are true at different instants fit together into a system, so that although the total course of history will be differently described at different times, the description at one time will determine what the descriptions at other times will be. [Prior and Fine 1977, p. 38]

As we saw, this is what motivates negative presentists such as Bourne to construct just such a 'system' of (ersatz) moments with appropriate links between them (see p. 227 above). But, as should be clear from our above worries, Prior here states something that is impossible to state *within* tense logic. As Sebastian Rödl observes, considering an intellect that would only think thoughts of the forms that tense logic describes:

> Such an intellect *is* a totality of tense-logical contents, wherefore the system of these totalities is not accessible to him. One instance follows the other, and with it one totality takes the place of the previous one. . . . The meaning of a tense-logical sentence resides in the conditions under which "is-present true" applies to it, and these are different at different times. Hence, as time passes, the meaning of all tense-logical sentences shifts. . . . Since the meaning of *all* of his sentences shifts, he has no means to *say* that it shifts. [Rödl 2012/2005, p. 106]

Leaving the tense-logical intellect aside, we may say that a tense-logical *reality* is a totality of tense-logical facts, wherefore the supposed system of these facts cannot be part of reality. *That* yesterday's p is the same as today's $\mathbf{P}p$ is something that cannot be included in reality as a tense-logical fact—for at least three reasons. Firstly because it will then itself be in need of being put into the system by way of further facts of cross-temporal fact-sameness, leading to a regress; secondly because p and $\mathbf{P}p$ are simply *not* the same; and thirdly because the fact-sameness holds between facts in different tense-logical systems, not within one such system. Of course, there is no problem if p and $\mathbf{P}p$ are taken not as irreducibly tensed facts but rather as two perspectivally different versions of an underlying *atemporal* fact (say, p-at-t)—but that amounts to a B-theoretic underpinning of the tensed facts.

Rödl is thus making the point that, *if* we take Prior seriously when he says that tense logic is 'metaphysically fundamental', *then* there cannot be, within the tensed facts themselves, a system into which the 'totalities of tensed propositions which are true at different instants fit together', as Prior claims.[301] It seems, then, that taking

[301] To be sure, a more sympathetic reading of Prior would take him to be saying that *what tense logic aims to capture* is metaphysically fundamental, in which case we are in full agreement, as we said earlier—what tense logic aims to capture is (an aspect of) the second conceptual gear.

tense logic to be metaphysically fundamental forces an understanding of tensed facts upon us that is incapable of tracing them across time. In that sense, tense and time are separated. Fine's analysis of A-theory illustrates the point quite faithfully—in particular his non-standard A-theories: how the relativist's many realities and the fragmentalist's many fragments fit together into a system is never discussed by him.

Because of this separation of tense from time, the problem Fine himself raises for the standard A-theorist—that his 'now' may, for all he says, be frozen—becomes an urgent one. If tense is not thus separated from time, there may well be no such possibility of frozenness. That would be the case if the applicability of the tensed form of predication *requires* the passage of time.[302]

This separation of tense from time, moreover, betrays an underlying deeper problem. For reconsider the claim that the basic stock of statements for tense logic consists of present-tense statements. We saw, in the case of negative presentism, that such a basic stock of purportedly present-tensed statements threatens to collapse into a basic stock of merely atemporal statements. What is it, then, we may ask, that makes the basic stock of statements for tense logic contain present-tense statements, instead of merely atemporal statements? Rödl puts the problem as follows:

> [If] a present tense formula is elementary, then the contrast to the corresponding past tense formula is not part of its [content], and without this contrast, "present" has no temporal meaning. [Rödl 2012/2005, p. 107]

One might think that the tense operators **F** and **P** provide the required contrast. The problem, however, is that they are supposed to operate on the basic statements, whose content cannot therefore be dependent on whatever it is these operators do. Compare the modal case: on one very popular (but first-gear) reading, the modal operators $\Box$ and $\Diamond$ are supposed to apply to independently meaningful, a-modal propositions (and the operators themselves are then treated as quantifiers over possible worlds). Similarly, then, **F** and **P** are likely to be understood as tense operators that apply to independently understood, a-temporal propositions (and the operators themselves are then treated as quantifiers over instants). Again, we observe that times are being treated much like possible worlds.

It seems, thus, that tense logic's conception of basic tensed statements threatens to collapse into a conception of *atemporal* statements after all, and leaves us largely in the dark as to what it is the so-called 'tense'-operators are doing. Perhaps the only sense that can be made of them is in terms of ersatz times à la Bourne. Taking tense

[302] Recall our brief discussion of Pooley's take on Prior's understanding of the passage of time above (§6.2.3, p. 229).

seriously in the Priorian way we have been considering (i.e., by taking tense logic metaphysically seriously) turns out to be very questionable: despite the seemingly obvious presence of tense in the facts that are said to constitute reality, in the end the forms of these facts still depend fundamentally on first-gear feature predication.

What we discover, then, is that it is possible to construct a first-gear simulation of tense. The recipe goes roughly as follows. B-theorists say that temporal truths should be understood as being of the form *p*-at-*t*, where *p* involves atemporal, first-gear feature predication, and *t* denotes the temporal location on the time line at which the given truth finds its home. Now, supposing that *t* is the present moment, that same truth can be presented simply as *p*, where the time reference is implicit. Let us call truths of this form 'present-tensed'. We now add tense operators which shift the implicit time reference forwards or backwards over the time line (either over the eternalist's real time line or over Bourne's ersatz time line)—and *voilà*, we have a first-gear simulation of tense. Saying of *such* an understanding of tensed truths that it is 'metaphysically fundamental' doesn't make much difference with being a B-theorist. It only complicates matters, as we have seen in our discussion of Fine's versions of A-theory.

This result should not be too surprising. Tense logic, after all, is formally just an extension of predicate logic with modal operators. And predicate logic, as we have repeatedly pointed out before, only knows of first-gear feature predication. Of course, this is not true of propositional tense logic, since propositional logic is silent on the form of predication involved. There are, then, two options. On the first option, tense logic is an attempt to formalize an aspect of the second conceptual gear, in which case we have no quarrel with it except to point out that it is imperfect. On the second option, however, tense logic is stipulated to capture *what tense is*: then, for tense operators to be applicable to a given stock of elementary propositions, these propositions have to be understood independently of the tense operators, which implies that they cannot be understood as being present tensed—unless 'present tensed' is taken to mean 'atemporal'.

To conclude: it is possible to be an A-theorist and even a presentist, and talk *as if* one were employing the second-gear form of predication, while all the same still relying solely on first-gear feature predication augmented by a set of 'tense operators'. Now, given that first-gear predication applies to first-gear entities, we should find that first-gear versions of A-theory and presentism construe their tensed facts as being facts about such first-gear entities.

That is indeed what we find. Consider, for illustration, the non-standard A-

theories. An object that persists through a stretch of time can, on these views, be understood as a perduring thing, in which case it consists of a series of first-gear entities that are its temporal parts. These temporal parts, however, are involved in a much broader range of truths than on the usual versions of perdurantism, for the truths about them include what was and what will be true of them (in the sense of the Priorian tense operators), and even these are replaced by fresh ones for every time slice, as we saw—every moment has its own tensed facts. Alternatively, such a persisting object can be understood as an enduring thing, in which case the very same thing exists in multiple realities or fragments. It has one set of properties relative to one reality (or as it occurs in one fragment), another set in another. Again, the facts about it relative to one reality (or as it occurs in one fragment) include what was and what will be true of it (in the sense of the tense operators). But it is clear that what we have, here, is a form of time-relative endurantism, where the relativity is to be understood as not internal to the facts (as in the version we discussed earlier), but rather as external to those facts. That is, the relativity resides not in the facts themselves, but in their belonging to one or another reality (or fragment). As we noticed above (§6.2.1), such an endurantist view indeed treats objects as first-gear entities.[303]

Recently, Correia and Rosenkranz have proposed an alternative non-standard A-theory (in Fine's sense) that does, in effect, take some of the thoughts we just formulated seriously. In their critique of Fine's analysis, they accuse him of tacitly endorsing a principle they call 'Fixed-Content':

> If at time t, tensed fact f both constitutes reality and is the fact that p, then for every time $u, \ldots,$ at u, f is the fact that p. [Correia and Rosenkranz 2012, p. 310; slightly simplified]

In short, Fixed-Content says that facts retain their (tensed) contents over time. Hence, if fact f *now* is the fact that you are reading, then f *always* is the fact that you are reading. Now, as Correia and Rosenkranz rightly remark:

> Fixed-Content is a principle about the diachronic identity of tensed facts. To our knowledge, the question to which this principle answers has hardly ever been raised, and accordingly the principle has hardly ever been explicitly formulated. [ibid, p. 313]

[303] With respect to versions of standard A-theory, it is more difficult to make our point. In short, we may say that a persisting object is either only its present time-slice (in the case of negative presentism), in which case it is an atemporal entity, or it exists at multiple times (in the case of eternalist standard A-theory—moving spotlight theory), in which case it is again like the time-relative endurantist's atemporal objects.

There is, of course, a very good reason why nobody has ever raised any question about the *diachronic* identity of facts: the facts don't change, hence no temporal differentiation as to their individuation is required (i.e., no distinction between diachronic and synchronic individuation). But Correia and Rosenkranz think that if we wish to endorse tensed facts, we should give up on this natural thought, and endorse an alternative to Fixed-Content which they call 'Shifty-Content':

> If at time t, tensed fact f both constitutes reality and is the fact that p, then for every time u, if u is n days from t . . . , then at u, f is the fact that $(-n)$ days from the present, p. [Correia and Rosenkranz 2012, p. 314]

Translated into somewhat less technical terms, the thought is that facts undergo qualitative change in their content over time while remaining the same. Their diachronic identity is not determined by the mere *identity* of their content, but rather by the appropriate connection between the temporal distance involved ($-n$ days from the present) and the rest of the content (p), according to the straightforward rule embodied in Shifty-Content. According to Correia and Rosenkranz, these considerations invalidate McTaggart's argument as construed by Fine: the four principles *A-Theory*, *Neutrality*, *Absolutism* and *Coherence* only form an inconsistent tetrad if Fixed-Content is assumed. They dub their version of A-theory, which accepts all four principles but rejects Fixed-Content in favor of Shifty-Content *dynamic absolutism*. Dynamic absolutism in effect attempts to build the system of tense-logical contents which Fine, following Prior, fails to account for, into the persistence conditions of facts.

Now, a moment's reflection reveals that dynamic absolutism does not yield the kind of view we are looking for. The reason is that its understanding of temporal facts still assigns tensed truths a location on the time line. Indeed, that is inevitable if one holds on to *Neutrality*. Moreover, and relatedly, the view rests on first-gear feature predication, for the fact that, e.g., you were having breakfast yesterday morning, decomposes in two components: a relative time reference, and a content to which this time reference applies. That content is supposed to be 'present tensed', but, as with negative presentism, such present-tensed truths are very easily read atemporally. Indeed, the form of these temporal facts resembles a B-theoretic understanding of temporal truth even more than in the cases we have seen so far. One can arrive at something very much like dynamic absolutism if one takes standard eternalist B-theory and then adds copies of all the facts of every moment to every other moment in such a way that the correct temporal distance is included in these facts (thus, if p-at-t is a B-fact, then also p-$(-n)$-days-ago-at-$(t + n)$ for every n). What is gained, in comparison to Fine's versions of non-standard A-theory, is that no fragmentation or

relativization is required.

Be that as it may, we may extract from dynamic absolutism a crucial ingredient for the kind of A-theory we are looking for. The feature of dynamic absolutism that interests us is that it resolves the complete isolation of tensed facts at their proper moments, which we observed in the Priorian or Finean conception of temporal truth. It is, thus, one way of explicating what Prior seems to have had in mind: that different contents (say, p and $\mathbf{P}p$) may, at different times, somehow latch onto the same fact. A disadvantage of the view is, however, that it does not allow for *truths* to be similarly absolute, because truth applies to the *contents* of the facts: what *was* true yesterday *is* no longer true now, because the fact that had the appropriate content yesterday now has a different content. We find this to be a disadvantage because, as we remarked earlier, on our view facts are identical with truths (and, more generally, because the facts don't change). We propose, therefore, to extract only the useful bit from dynamic absolutism. Instead of saying that the *content* of the fact changes over time, we claim that only the *sentence* we may use to pick out the relevant content changes over time. That is, we agree with Rödl, who writes:

> Prior mistakenly believes that temporal statements are true relative to a time because he does not distinguish between sentence and statement. We can say of a situational *sentence* that it is true relative to a time. That means that, using it at that time, one makes a *statement* that is *timelessly true*. In this sense, we can say that a past tense sentence is true if and only if the corresponding present tense sentence was true. This describes how situational sentences are linked to one another so that using them we can, in virtue of the time and across time, make the *same* timelessly true statement at different times. [Rödl 2012/2005, pp. 103–4]

The order of 'situational' (i.e., tensed) *sentences* which Rödl here points out is just what Correia and Rosenkranz are devising as the *contents* of their dynamic facts. We reject such inconstancy in the contents of facts—so one might ask, then, what the content of the facts *is*, if not what Correia and Rosenkranz propose? Are we not obliged to provide what the supposedly absolute contents of a fact look like *in a neutral manner*, such that it can then be seen how different tensed sentences can at different times express *that* content?

Here we have reached the crux of the matter. By asking for such a 'neutral' presentation of the content, one is in effect asking for an *atemporal* presentation, that is, a presentation from a perspective outside of time. The B-theorist will give his preferred tenseless account of such a content, and the various versions of A-theory we have seen, including dynamic absolutism and negative presentism, attempt to avoid the resulting anti-realism about tense *by providing an alternative, tensed account*

of such a content instead, which they borrow from Prior's tense logic. The point is, now, that as soon as we accept the challenge to provide such a 'neutral', i.e., atemporal account of the contents of temporal facts, we have accepted, without even noticing it, the notorious eternalist picture of the time line as providing locations for the facts to be true 'at'. Whether or not we incorporate tense into our preferred view of what such facts look like then becomes irrelevant. What matters is whether or not the eternalist picture is accepted. As we have seen, all of the versions of A-theory discussed so far endorse that eternalist picture. Even sticking to tensed facts in the way Prior asserted to be 'metaphysically fundamental', that is, as having essentially a tense-logical form, is not enough to prevent the eternalist picture from being endorsed. In fact, it complicates matters greatly by introducing a relative notion of truth which we simply do not need. It is surprising to what extent it is taken for granted that A-theorists have to endorse a Priorian, time-relative notion of truth.[304]

Thus, we refuse, following Rödl, to provide such an atemporal gloss on the contents of facts. Instead, we may say that the contents of, e.g., the fact that you are reading is just that: *that you are reading*. And tomorrow we can *say* the same thing by saying *that you were reading yesterday*. The only way of expressing temporal truths is *from within time*; as soon as we silently introduce an (eternalist) standpoint outside of time, such that we may first collect the relevant truths and then locate them on the time line, we in effect lose sight of time.[305]

Perhaps a comparison can help bring out what we intend. Suppose that the future is open, such that, in particular, it is not yet settled exactly which objects will come into being. Suppose, furthermore, that Socrates was such an object—many millennia ago, Socrates did not exist, and it was unsettled whether anything like him would come about. Now, assuming that sets exist, as first-gear entities, they exist timelessly. So Socrates's singleton—the set having Socrates as its only member—is such an entity. Yet it came into being when he came into being. Still, there is no question as to the persistence of sets: sets are simply not 'in time', even if their members are. After Socrates died the set still existed, for its existence is independent of time. Now, we can only think about that particular set insofar as we are able to think about Socrates,

[304] It is fitting here to refer to Gareth Evans's interesting paper *Does Tense Logic Rest Upon a Mistake?* [1985b], in which he surveys several ways of understanding the relativity of truth in Prior's tense-logical system (though in ways that do not touch upon the issues we are interested in here). We derived the title of the present section from the title of his paper.

[305] For the philosophy of language, this means that tense is not an optional feature of languages, but is a requirement for any language insofar as it is supposed to be able to express temporal statements. Which is not to say, of course, that this feature can only be realized in one way—there are, of course, various different systems, based on grammar or on lexicon or on both, by which a language can fulfill this requirement. [See also Rödl 2012/2005, pp. 111–2].

given that the set is individuated by Socrates himself. That is, we can only consider Socrates's singleton insofar as we can think about an object that ceased to exist long ago. Socrates recedes ever more into the past, but his singleton does not.

The same holds, we submit, for temporal truths (or facts), and this, by the way, shows that the so-called truthmaker arguments against presentism are confused. Consider the truth that Socrates was wise: what makes it true? Of course, given that Socrates is long gone, Socrates cannot *make* it true (in the present tense). But we may say that Socrates *made* it true, by becoming wise, and, since truths are independent of time just like sets are, that truth is available as we speak, so to say. Indeed, insisting on a *presently existing* truth-maker for that truth amounts to a misunderstanding of temporal truth, for it is precisely typical of temporal truths that one needs differently tensed sentences to express them as time progresses.

If you like, we could take these sketchy remarks to constitute our own A-theoretic version of truth-maker theory. Temporal truths are made true *in time*, and in a very literal sense, by the things they are about: Socrates made it true that Socrates was wise, as said, and you are now making it true that you are (still) reading this chapter. There is nothing special about this activity of truth-making, of course—Socrates made true his being wise simply by becoming wise, and you are making true your reading this chapter simply by reading. It is the very nature of temporal truth that takes care of the rest: temporal truth is absolute. We do not need the (atemporal) existence of past things, nor do we need any ersatz truthmakers, for past truths to be available at the present moment. Once made true, their availability is guaranteed.[306] Thus, properly understood, positive presentism and its accompanying version of A-theory do not give rise to the problems that notoriously plague other versions of presentism. Hence these 'problems' are trivially 'solved' once we understand what temporal truth amounts to, on the second conceptual gear.

What we have arrived at, now, is the view that a proper A-theory will not endeavour to provide an account of time from an eternalist starting point—on pain of collapsing into mere insistence on tensed facts understood, ultimately, in a first-gear way. A proper A-theory will take tense seriously in the sense that it takes temporal facts or truths to be both *inherently* temporal and *existentially* temporal, as we might say: inherently temporal by featuring a tensed form of predication, and existentially

[306] Our literal version of truth-making is reminiscent of the so-called 'growing block' theory of time. The growing block view says that reality comprises all things past and present, but not future, and that 'fresh slices of existence' are added as time progresses [see Broad 1923, esp. p. 66]. Substituting facts or truths for things, our own view of timeless yet temporal truths drops out. (Contemporary defenders of the growing block view include Adams [1986, esp. p. 322] and Forrest [2006]; Earman [2008] constructs a relativistic version of the view—see also §6.2.6 below. For recent criticism, see Merricks [2006].)

temporal by themselves originating 'in time'. Even the contents of temporal truths have to be presented from within time—that is, from the present moment—without thereby distorting their absoluteness. That is what we propose the phrase 'taking tense seriously' to amount to. It is different from just having facts whose contents are expressed using the linguistic device of the tenses, which is what Fine, Prior, and others at times seem to have assumed.

We have made progress. We have arrived at a view on tensed facts which shows temporal truths to be, in a way, temporally absolute yet not atemporal. This form of A-theory does what Merricks wanted it to do (see the quotes at the end of §6.2.1): it prevents the perplexing, contradictory picture of change from emerging. For, as we saw, change only seems to result in a contradiction if we translate the tensed description of the change-involving situation into an atemporal form. If the banana was green yesterday but is now yellow, we may say that the truth we could express yesterday by saying 'the banana is green' simply *is* (timelessly) the truth we can now express by saying 'the banana was green', which is not in contradiction with its present yellowness. We simply observe that the banana has produced more truths since it was green, one of which is its successfully having turned yellow.

The tensed, second-gear form of predication *is* time, rather than being useful for a *reduction* of time. Instead of providing an essentially atemporal picture of reality within which we can then go on to single out what we call 'time' (or 'tense' or 'persistence'), we now provide an essentially temporal picture of reality—that is, a picture of reality as itself persisting through time, or, in other words, a presentist picture. No materials for the reduction of change or persistence is therefore available, since the very existence of reality as temporal depends on these matters. What Pooley took Prior's tense logic to achieve (see p. 229 above), can thus only be achieved if we amend Prior's take on tensed truth in the ways we have discussed.

6.2.5 Positive Presentist Persistence

We have noticed that large parts of the debates between endurantists and perdurantists, between A-theorists and B-theorists, and between presentists and eternalists, can be understood as battles within a first-gear arena. Therefore, our truly second-gear versions of presentism and of A-theory are bound to be unacceptable to participants in those debates, for they consist largely in refusing to conform to the way in which those debates are very often construed. The best a second-gear thinker can do is try to point out just where his statements are being illegitimately assimilated into

the kind of first-gear conceptual framework he rejects—and that is, indeed, in large part what we have attempted to do in our investigations so far.

In this section we return to the topic of persistence, which we explored in §6.2.1 above. There, we finished by concluding that a proper version of endurantism requires A-theory, and in the subsequent subsections we worked our way towards a proper, second-gear understanding of A-theory and presentism. Let us see what we can say about persistence on this basis.

We start with another instance of what we just noted: that the issue of perdurantism versus endurantism is in large part understood as a purely first-gear issue. Hawley, for instance, claims that endurantism is unformulable once an atemporal understanding of persisting objects is rejected—that is, once we reject first-gear thought. Let us have a look at her reasons for drawing this bold conclusion:

> According to endurance theorists, it is one and the same object—the banana—which is first green all over then yellow all over. . . . [T]o endure, an object must exist at more than one time. Now, if we give up atemporal talk, how can we assert that the green banana and the yellow banana are identical? . . . We might say that the yellow banana is [now] identical with the green banana, which is to say that the yellow banana is [now] identical with something which was the green banana. [Hawley 2002, p. 31, slightly contracted]

So far, so good—what Hawley here proposes can be understood to represent the position we are defending. However, she takes this proposal to be defective:

> But it is common ground between many accounts of persistence that the yellow banana used to be the green banana. Both endurance and perdurance theorists will accept that the yellow banana used to be the green banana, although the theories can give different atemporal descriptions of the underlying reality so long as atemporal talk is permitted. . . .
> If we are only permitted to talk about how things are [now], then the most we can do in speaking of persistence is to speak of the histories and futures of objects . . . . We cannot assert or deny claims of identity between objects existing at different times, and thus endurance theory is unformulable. [ibid, pp. 31–2, slightly contracted]

What Hawley here calls 'common ground' is not really common ground at all. For as soon as she says that 'different atemporal descriptions of the underlying reality' can be given, she makes clear that she demands the participants in the debate over persistence to adopt a basically first-gear conceptual framework. *That* is what she takes the common ground to be, and that is precisely what a second-gear thinker will reject.

Hawley thus, rather ironically, pictures the debate on persistence to be a debate about the correct first-gear simulation of persistence:

> Those who adopt an irreducibly tensed view of the temporal world, and do not accept that a tenseless description can ever be even partially adequate need have no truck with the debate between endurance and perdurance theories. [Hawley 2002, p. 34]
> The task of this book is to explore the debate about persistence, not to convince those who believe that there should be no such debate .... [ibid, p. 31]

The second-gear thinker should, of course, object to such a kidnapping of one of the fundamental categories of the second conceptual gear: persistence. And that is exactly what we are now doing.[307]

Hawley's way of putting the controversy between endurantists and perdurantists, namely, over the way in which questions of cross-temporal identity are settled, illustrates an important first-gear ingredient in her thinking about persistence: persistence is to be understood in terms of first-gear existence. Suppose that a persists from t_1 to t_2. At time t_1, there exists (atemporally) a certain object, a_1, and at time t_2 there exists (atemporally) a certain object, a_2. The perdurantist says that they are not identical, that $a_1 \neq a_2$, but parts (atemporally) of a temporally extended perduring thing, a, while the endurantist says that they are identical, that $a_1 = a_2$, and hence that $a = a_1$ and $a = a_2$.

By contrast, the second-gear thinker will reject the very assumption that persistence presupposes atemporal existence in this way. For physical objects, to exist *is* to persist. Of course, there can be questions of 'cross-temporal' identity: is the planet I saw in the morning sky the same as the planet I now see in the evening sky? But such questions are not much different from other questions of identity, such as: is the man I see in this mirror the same as the man I see in that mirror? (supposing that we are standing in a confusing house of mirrors.) Both questions are about persisting things (note the tenses); they are not about situations in which atemporal, first-gear existence may or may not amount to persistence—it never does.

Nothing can be learned about persistence by starting in the way first-gear thinkers like Lewis, Sider and Hawley do. Nor do these philosophers leave any room for doubt about their first-gear proclivities. Compare Lewis's presentation of endurantism:

> The hypothesis of endurance ... [says that] things have no temporal parts. Rather, a persisting thing is multiply located in time: the whole of it is at one time and also at another. Endurance calls to mind ... the power of spatial bilocation traditionally ascribed to saints. If a bilocated saint is wholly in Rome and wholly in Byzantium, and if in Rome he is bent and in Byzantium he is straight, then we

[307] Another example is Lewis, whose conception of presentism coincides with what we called negative presentism—on such a view, he rightly claims, there is no persistence at all. Hence he takes the debate over persistence, in effect, to presuppose eternalism (or 'tenseless description'). See the quote we presented on p. 227 above.

> have a problem of local intrinsics that exactly parallels the problem of temporary intrinsics for an enduring thing . . . . [Lewis 2002, pp. 2–3]

It is obvious, then, that the version of endurantism embodied in our second-gear understanding of time will be rejected by many participants in the contemporary debate over persistence—especially by perdurantists, but also by endurantists who (tacitly or explicitly) accept a first-gear construal of the issue.

Should we then perhaps drop endurantism, and say that we subscribe to a different, second-gear understanding of persistence, for which a new -ism needs to be introduced? No: typical endurantist statements, such as that persisting things are 'wholly present whenever they exist', that they 'have no temporal parts', and that they 'have no temporal extension', play an important role in our second-gear understanding of persistence. Indeed, those defending endurantism in one of its first-gear guises are often *in fact* attempting (but failing) to capture these slogans because of their second-gear import. That is what makes the debate so confused.

To see what the second-gear import of such endurantist slogans consists in, consider the question whether reality can be temporal, that is, such that the second conceptual gear applies, without there being any persistence over time. One option would be to take reality to be inhabited by instantaneous objects, things that are destroyed right when they are created, such that they do not exist longer than an instant. Presumably, such objects are a limiting case of persisting objects. On this assumption, our supposedly temporal reality collapses into a series of first-gear realities—much like the perdurantist's view. To see why, first consider whether or not there is any process going on in this reality. To have a process, there has to be an object that is engaged in that process. So, suppose that object a is doing V. As we observed, this implies that *a has done V*: processes take time (see §4.3, especially argument [G2.6] on p. 146). But that a has done V can only be true if a has existed longer than an instant. So there are no processes. But if there are no processes, there is no change, and without change, there is no time. Hence persisting objects are not an optional ingredient for a temporal reality: only by having persisting objects can there be change. In short, change requires persistence.[308]

What is crucial for the second conceptual gear is, thus, that there are persisting things *that remain the same over time*, i.e., that are wholly present whenever they exist. That is, persistence is understood along endurantist lines, the notions of temporal

[308] Conversely, one could argue that persistence requires change: a supposedly temporal reality which is 'so boring that the same tensed facts hold at every single time' [Fine 2005d, p. 272] collapses into a merely spatial, i.e., first-gear reality that does not involve time at all. Admittedly, this does not yet exclude the possibility (if it is a possibility) of a global 'freeze' in time—as envisaged by Shoemaker [1969].

parts and temporal extension are to be rejected. Instead of inventing a new -ism for second-gear persistence, we should claim that a first-gear assimilation of endurantism (and thereby of the entire debate over persistence) muddies the conceptual waters by obliviating what the debate on persistence (like the debate on time in general) is really about: first-gear reductive thinking versus second-gear non-reductive thinking.

6.2.6 The Open Future and Relativity Theory

We would like to finish our exploration of time as a basic category of second-gear thought with a brief reflection on two topics largely ignored so far: the idea of an open future, and the argument against presentism based on special relativity. We start with the former—in the words of cosmologist George Ellis:

> Things could have been different, but second by second, one specific evolutionary history out of all the possibilities is chosen, takes place, and gets cast in stone. [G. F. R. Ellis 2006, pp. 1812–3]

As we observed earlier (in §4.3, see esp. p. 140), the second conceptual gear does not tell us anything about whether or not the future is open, but it does tell us what it is for the future to be open or not. As we will see in the next chapter, the third conceptual gear presupposes an open future in just the sense Ellis tries to capture (see §7.2). We observed, during our presentation of the versions of presentism, A-theory and endurantism that the second conceptual gear gives rise to, that there is a strong tendency to assimilate each and every fundamental second-gear concept into a basically first-gear setting. The notion of an open (or closed) future is no exception: a purely first-gear understanding of this notion divorces the issue of determinism and indeterminism entirely from its proper conceptual anchor—the second conceptual gear.

A clear illustration of what we have in mind here can be found in the following statement by J.J.C. Smart:

> We can now see also that the view of the world as a space-time manifold no more implies determinism than it does the fatalistic view that the future 'is already laid up'. It is compatible both with determinism and with indeterminism, i.e. both with the view that earlier time slices of the universe are determinately related by laws of nature to later time slices and with the view that they are not so related. [Smart 1963, pp. 141–2]

It is obvious that Smart's starting point is a first-gear understanding of reality (as a 'space-time manifold'), in terms of which he understands such concepts as laws of nature and, consequently, determinism and indeterminism.

By contrast, a second-gear understanding of determinism and indeterminism hinges on the kinds of powers that things may have, as we already suggested in §4.3. There, we distinguished between deterministic powers on the one hand, and indecisive and spontaneous powers on the other hand. Deterministic powers, so the thought goes, are such as to result in only one specific kind of causal process when the right circumstances are present. If there are only deterministic powers, the way things will continue from now is entirely determined: that is determinism, as a second-gear notion. Notice that determinism, thus understood, presupposes the presentist/endurantist/A-theoretic picture we have been working towards. If, on the other hand, there are spontaneous powers, which may or may not manifest irrespective of the conditions, or if there are indecisive powers, which may respond with a range of different manifestations to the very same conditions, the future is open: that is indeterminism, as a second-gear notion. Again, indeterminism, thus understood, presupposes the presentist/endurantist/A-theoretic picture we outlined above. In both cases, there is no talk of 'earlier time slices' and 'later time slices' that are or are not deterministically related by 'laws of nature'. There is only talk of reality as it is now, and of how this reality may unfold into the future.[309]

Given this second-gear understanding of determinism and indeterminism, it is clear that the charges which Smart so carelessly brushes aside—that his view of the world as a space-time manifold implies determinism and the fatalistic view that the future 'is already laid up'—cannot be dismissed by appealing to the first-gear surrogates of determinism and indeterminism he suggests. Admittedly, Smart's picture is not deterministic in the second-gear sense, given that it is not even a second-gear picture. But insofar as he alleges that picture to adequately capture the reality we live in, it indeed follows that the future 'is already laid up': that is exactly what the space-time manifold picture amounts to.[310] As before, we have to conclude that the real controversy is between first- and second-gear thought: Smart refuses to take anything like our second-gear understanding of determinism and indeterminism seriously, and hence sees no good arguments against his view, while second-gear thinkers simply fail to find the very features of reality they care about—

[309] A rigorous formal explication of the second-gear picture we have in mind is being developed by Rumberg and Müller [2013] and Rumberg [n.d.]. Of course, formal systems are bound to be amenable to a purely first-gear interpretation, and that is true also in this case. However, the formal system Rumberg develops illustrates nicely how first-gear tools can be used to shed light on second-gear concepts, if that strong urge to reduce, that bedevils many analytic philosophers, is resisted.

[310] As Oliver Pooley puts it, the future is on such a view admittedly not *determined*, but still *determinate* [Pooley 2013, p. 337]. See McCall [1976] and E. Barnes and Cameron [2009, 2011] for defenses of the thesis that an open future requires indeterminateness.

time, persistence, causality, an open future—in a first-gear picture like Smart's.

In the philosophy of science, it is common to think of determinism and indeterminism as properties of theories: a given theory (construed as an axiomatic formal theory) is deterministic if, for instance, any two of its models that agree on what is true at one point in time agree on what is true at all times—or, alternatively, if given a complete description of one point in time, what is true at other points in time is deducible without remainder.[311] The world is deterministic, then, so the thought goes, if the fundamental theory of everything that our sciences (allegedly) aim at is deterministic in this sense.

A moment's reflection reveals that such an approach to understanding determinism and indeterminism is hopelessly equivocal between a first- and a second-gear reading—and that is an interesting result in itself. To bring this out, consider Russell's argument to the effect that determinism collapses into a triviality—which Earman lucidly summarizes as follows:

> [I]magine a very simple universe containing a single dimensionless particle, and suppose that the state of the particle at any instant t is specified by its position coordinates x_t, y_t, z_t. The motion of the particle through space can be as complicated as you like as long as it can occupy only one place at a time. Then, as a matter of mathematical fact, there must exist functions f_1, f_2, f_3. such that $x = f_1(t)$, $y = f_2(t)$ and $z = f_3(t)$. The example can be made more realistic by adding other particles and additional state variables, but the essential point remains the same. [Earman 1986, p. 11]

Earman goes on to quote, by way of conclusion, the following passage from Russell himself:

> It follows that, theoretically, the whole state of the material universe at time t must be capable of being exhibited as a function of t. Hence our universe will be deterministic in the sense defined above. But if this be true, no information is conveyed about the universe in stating that it is deterministic. [Russell 1953, p. 401]

The thought is, then, that no matter how things are, there will always be a theory which exhaustively describes everything past, present, and future given how things are at a particular moment—a deterministic theory, that is. Now, apart from the question how one might avoid such an apparent trivialization of the concept of determinism, it is interesting to consider how exactly the function (or theory) Russell

[311] The most important reference here is, of course, Earman [1986], though part of his argument is that such a model-theoretic or syntactic approach to determinism should be abandoned in favor of an approach that is closer to the form of actual physical theories—i.e., should revolve around questions concerning the existence and uniqueness of solutions to differential equations [see esp. p. 21].

had in mind works. As we noted earlier (in §4.3, p. 147), we may understand the applicability of such a function to reality in two different ways. We may take it to be applicable to reality conceived of as a four-dimensional, static block: then, we are adopting a first-gear stance, and we conceive of the particle from Earman's example as a first-gear entity which atemporally occupies different locations at different 'times'. But we may also understand the function to be applicable because the particle has a certain power which dictates how it moves, where the movement is understood as a second-gear process. In the first case, the 'movement' of the 'particle' is reduced to a curve through a four-dimensional manifold, in the second case that mathematical curve is derived from the concrete movement of the particle. In the first case, the curve as a whole is a brute fact, in the second case, the curve is explained by the powers of the particle. In the first case, then, determinism is conceived in the same way Smart conceived of it, while in the second case it is conceived of in our second-gear way.

For the second-gear metaphysician, then, Russell's observation is completely irrelevant: what matters is that we capture the powers and processes that are really there. Determinism is then no longer trivially true, nor is indeterminism then trivially false: the open future requires such non-trivial indeterminism, consisting in a range of real possibilities for the world to continue branching off from the present moment.

These real possibilities may, as is well known, be represented by use of branching models.[312] And here, again, we need to take care not to lapse into a first-gear understanding of such branching models. For as soon as one does, the result becomes very strange indeed—does reality itself branch? And how do we then know on which branch we will be the next minute, if all the future branches are equally real? Now, as Oliver Pooley writes:

> On the view we are exploring one *starts* with the intuition that ... the future is open. This is taken to mean that (i) there are several ways that things might happen, and (ii) nothing in reality singles out ... one of these possibilities as the way things will actually be. On the intended interpretation, branching-time structures are introduced as a means to represent these purported facts, and to see whether a consistent formal theory incorporating them can be devised. Against this background, one is simply not permitted to reinterpret the branches relative to some time as several equally real futures, rather than as several equally real possible ways that the single future might turn out. [Pooley 2013, p. 339]

In fact, this is just to repeat some of the points we have been making in the course of our discussion of A-theory (and presentism), but now in application to the idea of real future possibilities.

[312] See fn. 309 on p. 249 above.

Theories couched in formal or mathematical terms (be they functions of the kind Russell imagined or differential equations or branching models) do not tell, all by themselves, how they are to be understood. First- and second-gear metaphysician alike may subscribe to the same physical theory, mathematically speaking, or to the same model-theoretic representation, formally speaking, even though the former understands it as a summary of the given four-dimensional scenery while the latter understands it to concisely present the ways in which certain fundamental powers and processes work and interact. We may use this insight to briefly touch upon that one big issue we have neglected so far: the theories of relativity. For it is obvious that not every theory is amenable to such a second-gear reading (although, interestingly, first-gear readings are always available, though often bizarre). The question, then, is whether the theories of special and general relativity can be endorsed by a second-gear thinker, and if not, what this implies. We do not intend to settle these difficult questions, but merely explore what is required to arrive at an answer.

It is often claimed that presentism, and A-theory in general, cannot be true precisely because it is incompatible with the theories of relativity. These theories, after all, cast doubt on the cogency of an absolute notion of simultaneity, yet such a notion seems to be central to presentism and A-theory.[313] Now, obviously, the thought that space and time cannot be separated as cleanly as classical presentism and A-theory require does pose a challenge. The challenge is to come up with an interpretation of the theories of special and/or general relativity—which are based, after all, on a purely mathematical core—that is compatible with the second conceptual gear. Obviously, it is much easier to understand, e.g., special relativity's Minkowski space-time as a first-gear four-dimensional manifold, and endorse the eternalist picture of reality that this gives rise to. Yet advances have been made in very recent research on these matters that at least make clear that such a first-gear picture is not mandatory for second-gear thinkers who wish to take relativity seriously.[314]

The most promising work in this area seems to stem from attempts to account for the open future in a relativistic setting. Nuel Belnap, for instance, introduced the idea of 'branching space-times', which hopes to generalize the classical branching times

[313] See, e.g., Saunders [2002], Peterson and Silberstein [2010], and Wüthrich [2010] for arguments that relativity refutes presentism, and Hinchliff [2000] and Friebe [2012a,b] that it does not—Friebe argues, moreover, that B-theory is threatened by relativity just as much as A-theory. Philosophically interesting discussions of the theories of relativity, that bear on the questions at hand, can also be found in the essays contained in Dieks [2006, 2008].

[314] Another route that is available for the second-gear thinker, which we should not wish to reject right out of hand, consists in pointing out that both special and general relativity are, after all, theories which are known to be strictly false, so that the challenge we formulated need not even be accepted.

approach to formally representing the open future to the relativistic case. Now, here as in the classical case, we should take care not to understand the proposal along first-gear lines (and we should note that doing so results in a first-gear theory of a branching reality not unlike the Everettian many-worlds interpretation of quantum mechanics).[315] Rejecting such adventurous metaphysical extravagance, it is clear that attempts such as Belnap's only make sense when the mentioned challenge is met: the branching spatiotemporal manifold must somehow represent genuine future possibilities in a relativistic setting—not eternalist actualities (recall Pooley's remark quoted above). In his recent discussion of the prospects for such an understanding of a relativistic branching-based formal system, Pooley writes:

> it is important to distinguish formal models from how these models are supposed to express the central commitments of the defender of real passage [i.e., the second-gear concept of time]. In these terms, generalisation of A-theoretic views to relativity involves two stages: (1) identifying relativistic analogues of the classical models; and (2) explaining how such models articulate a non-B-theoretic view of time. The first step is relatively straightforward, and generically involves replacing a structure of totally ordered elements (intended, for example, to represent the tensed facts that hold as of some time) with a structure whose elements are only partially ordered. The harder task is to generalise to relativity the kind of stories that can be told about the pre-relativistic models. [Pooley 2013, p. 323]

Of course, we should add that avoiding B-theory is not sufficient, as we have argued above—but if we keep in mind the lessons we learned in reading this quote, we do arrive at a reasonably clear statement of what the project of understanding relativity from within the second conceptual gear involves.[316]

Pooley himself proposes an understanding of relativistic A-theory which is close to Kit Fine's non-standard A-theories—recall our discussion of Pooley's generalization of the way presentist A-theory accounts for the passage of time to non-standard A-theories in §6.2.3, p. 229 above. That is, he suggests that endorsing special relativity within an A-theoretic setting should preferably include the rejection of an absolute 'now' [see ibid, §VII].[317] Given our results so far, it is clear that such a route is not available for us: positive presentism is incompatible with non-standard A-theory

[315] For branching space-time, see Belnap [1992, 2012]. For Everett's interpretation of quantum mechanics, see, e.g., Everett [1957] and Albert and Loewer [1988].

[316] This project is being taken seriously by prominent researchers in several related areas; so, e.g., by cosmologist G. F. R. Ellis [2006] and philosophers of physics Norton [2010] and Weinert [2010].

[317] For similarly relativistic reasons, Fine takes his non-standard A-theories to be preferable over standard A-theory. Moreover, he takes the 'fragmentalist' version of non-standard A-theory to be preferable over the alternative, external relativist version of non-standard A-theory because it does not depend on an ontology of times, and allows for overlapping fragments [see Fine 2005d, §11]. Another defense of a form of relativistic presentism that appears to reject absolute simultaneity (in this case, by rejecting the transitivity of co-existence) can be found in Lowe [2013b].

because the latter kind of view rests on the idea, ultimately, that tensed truths are temporally located. It seems, therefore, that only approaches that add an absolute 'now', i.e., a metaphysically preferred foliation, to the relativistic space-time stand a chance to be acceptable from a second-gear point of view.[318]

Here our excursion on the topic of relativity and the second-gear concept of time ends. It seems to be largely an open question whether or not the two are compatible. We should stress, though, that if it turns out that they are not, this is reason to reexamine our explication of the second conceptual gear, but equally to doubt the adequacy of the theories of relativity themselves. Resorting to a first-gear understanding of time in response to such a situation would amount to giving up on our quest to understanding time. To repeat: the most important part of the challenge is to explicate how the mathematical or geometrical statements of the theories of relativity as well as of attempts to combine them with presentism or A-theory are to be understood.

6.3 Concluding Remarks

What emerges from our journey though the complex conceptual landscape of the contemporary philosophical discussions on causation and time is a picture that is, in a way, so familiar that it might strike one as a retreat into a kind of naive realism: we live in a (presentist) world inhabited by (enduring) objects which are constantly engaged in causal processes of various kinds. Instead of providing a metaphysical explanation of this reality *in more basic terms*, the second-gear metaphysician simply accepts these features of reality as basic in their own right, and remarks that the urge to find such metaphysical explanations in more basic terms arises solely out of a preoccupation with first-gear thought, born perhaps out of strong empiricist convictions.

The first conceptual gear is peculiar in that it does not in itself amount to a metaphysical view—except, perhaps, insofar as abstract objects are concerned. In our presentation of the conceptual gears in chapter 4, we saw that our description of the second and third conceptual gears differed in this respect from our presentation of the first gear. Consequently, in the previous chapter, we noticed that a first-gear metaphysical view needs to be constructed by providing analyses of basic higher-gear concepts in first-gear terms (see §5.1). In the case of the second conceptual

[318] This option has recently been elaborately defended by Zimmerman [2005, 2008, 2011]. For recent criticism, see Wüthrich [2013].

gear, then, no such task is required: the second conceptual gear in itself consists in a certain metaphysical picture—hence the disappointment on the part of those expecting clever bits of first-gear conceptual engineering. To be sure, many interesting results can be obtained by subjecting aspects of the second gear to thorough, first-gear scrutiny—but that is not the same as *reducing* the second conceptual gear to mere first-gear simulations.

We find that we understand reality in large part in second-gear terms (and in third-gear terms, on which more in the next chapter). On the Aristotelian picture, that means that we take the second conceptual gear metaphysically seriously. Although everything we have claimed for the second conceptual gear is of course open for critical discussion, we disregard the first-gear metaphysician's stance, which consists in *first* deciding that all of reality be intelligible in first-gear terms, and *then* requiring our second-gear concepts to be analyzable in conformity to this metaphysical dogma. That is typical of Modern-picture metaphysical realism, which we have dismissed in chapter 1.

It has been the aim of the present chapter, first, to bring out how deeply-rooted this basic first-gear stance can be, such that even those aiming to resist outright first-gear reductionism are frequently captivated by its compelling charm, and second, to clarify the commitments of the second conceptual gear by elucidating its two most basic categories: causality and time—or, put in terms of the second-gear form of predication, aspect and tense. Those two aims have been pursued in tandem: results pertaining to the first aim have resulted in clarifications of basic second-gear notions, while results pertaining to the second aim have contributed to unmasking certain views that might be taken to be second-gear as ultimately resting on a first-gear basis after all. We may conclude that attempts to mix second-gear elements into a first-gear setting invariably result in first-gear simulations, and thereby contribute to a widespread confusion within contemporary metaphysical debates: lots of apparently fundamental divides can be found that turn out to be merely internal disputes amongst first-gear thinkers. The fundamental divide, between accepting or not accepting the second conceptual gear, is thus easily lost from sight.

Looking forward now to the next chapter, we may notice that there is a similarly strong tendency to understand third-gear phenomena in purely second-gear terms. Physicalism, materialism, reductionism, naturalism—under such labels, vague and confusing though they invariably tend to be, many embrace the thought that, however complicated living beings are, their existence and persistence is ultimately to be explained in terms of second-gear physical components and the causal processes

by which these interact. Organisms are complex and adaptive mechanisms, so the thought goes. The structure of the controversy over what life is, is thus not unlike the clash between first-gear and second-gear thinkers: *first* it is decided that reality be intelligible in second-gear terms, and *then* it is required that the basic third-gear concepts be analyzable in conformity with this metaphysical dogma. Let us move on to the next chapter to see how this second-gear stance fails to capture what is distinctive of life.

* * *

Chapter 7

The Third Gear: Life

In this chapter, we turn to the most controversial but also the most interesting part of the argument of this essay: we move on to discuss life. We placed the concept of life at the heart of a distinct, third conceptual gear, and insofar we may be said to engage in a discussion of the 'logic of life'.[319] We presented a sketch of the conceptual gears that we discuss in this essay in chapter 4. By way of orientation, let us repeat our overall scheme of these three conceptual gears:

	First gear	**Second gear**	**Third gear**
form of predication:	tenseless	tensed/aspected	normative
applies to:	entity	object	organism
predicates involved:	feature	state/process	life-process
sortals involved:	type	natural kind	life-form
mode of being:	to exist	to persist	to live
engaged in:	—	process	life-process
characteristic:	geometry	causality/temporality	teleology/normativity
	Humeanism	*anti-Humeanism*	

The present chapter does not expand much on the sketch of the third conceptual gear presented in §4.4. Like the previous two chapters, it is devoted, rather, to a defense of the autonomy of the respective conceptual gear in the light of contemporary debates—this time in the philosophy of biology.

Our investigations into the first and second conceptual gears in previous chapters have shown that it is not easy to determine what exactly the basic conceptual

[319] Recall that we decided to not use the label 'logic' too readily; we will thus continue our habit of naming the investigation we are engaged in a conceptual one. See §3.4.

ingredients of the respective conceptual gears amount to, once one puts them in the context of the relevant contemporary metaphysical debates. And, relatedly, we discovered that it is equally difficult to find out against which conceptual background given philosophical views and debates, such as those on laws of nature (§5.2.2), causation (6.1), and time (6.2) are to be understood. The main lesson learned, for our overall project, is that adherence to what we called the Modern picture in chapter 1, and specifically to metaphysical realism, leads to philosophical disagreements of a typical nature: one specific conceptual gear (or an aspect thereof) is adopted as capturing what reality-itself is *really* like, and hence starts to operate as a metaphysical dogma to which concepts from higher conceptual gears have to conform. The result is a tendency to reductively analyze higher-gear concepts using lower-gear materials, to which those who appreciate the self-contained character of those higher conceptual frameworks then object. That is what the controversy between Humeans and anti-Humeans ultimately rests on, as we have seen.

On the Aristotelian picture, which we outlined in §1.3 and whose metametaphysical underpinnings we sketched and argued for in chapters 2 and 3, this tendency to form conflicting metaphysical dogmas does not arise: the diversity of conceptual gears mirrors—or, rather, *is*—a diversity in reality. That does not rule out critical reflection on every aspect of these conceptual gears—we may learn that we have made mistakes in what we take these conceptual gears to amount to. We may learn, for instance, that the division into gears we here propose should be replaced by a different one, that provides clearer insight into the conceptual and thereby into reality. The aim of the present essay is, first and foremost, to defend this kind of understanding of metaphysics—to defend the Aristotelian picture, that is. What we develop in Part II, to which this chapter belongs, is our best shot at a systematic development of a metaphysics on the basis of that picture.

So much for a sketch of the context of the present chapter on life. Part of what we have shown in earlier chapters is that first- and second-gear thought can be applicable to one and the same reality, without requiring reductive analyses of second-gear concepts in first-gear terms. In general, then, diversity in conceptual gears invites the following two questions: How can they all be applicable to one and the same reality? How does the noted diversity relate to the unity of reality? As we will see, considering these questions in the context of the relation between second- and third-gear thought opens up an entirely new dimension to them.

To start, we explore the conceptual autonomy of the third conceptual gear in §7.1 below, by following closely Michael Thompson's discussion of the 'real definition

of life'—as we said, we owe the very idea of distinguishing diverse conceptual gears to Thompson. We then move on, in §7.2, to the aforementioned questions, and explore the relationship between second- and third-gear phenomena in the context of a discussion of mechanisms and organisms. Thereby, we also provide an illustration of the noted tendency of Modern picture-based metaphysical realism to result in clashes between reductive analyses, in this case of third-gear concepts in terms of second-gear ones, and the self-contained character of what is thereby purportedly reduced. A further illustration of this point is developed in §7.3: there, we look at the work of Ruth Millikan, who uses basic categories of the third conceptual gear in her teleosemantic approach to questions in, e.g., the philosophy of mind and language, while also providing a reductive analysis of just these categories. Finally, in §7.4, we turn to the theory of evolution, in order to dispel two *prima facie* problematic clashes between that theory and our non-reductive, essentialist understanding of life. A brief conclusion summarizes our findings in this chapter (§7.5).

7.1 The Real Definition of Life

If living things do not comprise a (third-gear) category of their own, but rather form a mere subclass of concrete, physical (second-gear) objects generally, then the thought of providing a list of characteristics found only in living things, such as to single out that particular subclass, is a natural one. As Thompson remarks, '[t]he idea of such a list must have arisen during the vitalist-mechanist debates' [M. Thompson 2008, fn. 3 on p. 4]; he cites Moritz Schlick, who provides such a list in philosophical lectures of 1927 in the context of his discussion of precisely those debates.[320] Now, typically, biology textbooks, as well as other introductory texts concerning the science of the living, begin by providing a version of such a list. Take, for instance, the opening sentences of Wikipedia's 'Life' article:

> Life is a characteristic that distinguishes objects that have signaling and self-sustaining processes from those that do not, either because such functions have ceased (death), or else because they lack such functions and are classified as inanimate. ... Organisms undergo metabolism, maintain homeostasis, possess

[320] A translation can be found in Schlick [1949, pp. 73–4]. Schlick says he got his list from Wilhelm Roux, the founder of 'developmental mechanics', who was deeply involved in the vitalist-mechanist debate around 1890. In those years, Roux's experimentally developed mechanistic theory of embryological development was refuted by experiments of Hans Driesch, who argued, moreover, that his own results supported vitalism. See Mocek [1974]. We briefly touch upon the vitalist-mechanist debate below, in §7.2.1.

> a capacity to grow, respond to stimuli, reproduce and, through natural selection, adapt to their environment in successive generations.[321]

One may read these remarks as expressing the kind of (second-gear reductive) attitude we are highlighting: life is a 'characteristic' of physical 'objects' more generally, and involves metabolism, homeostasis, growth, responsiveness, reproduction, and adaptation. For such a list to be successful, its elements must be such as to be understood independently of life itself. Regarding lists that purport to serve such a reductive goal (explicitly or implicitly), Thompson suggests that the very project of providing such a list is a hopeless one, precisely because living things form a category of their own—a category, to be more precise, that finds its home within a broader range of categories of which it is a part: the third conceptual gear (see §4.4). Thompson writes:

> My suggestion will be that *every* candidate list-occupant must strike the sub-metaphysical Scylla of "DNA" or else sink into the tautological Charybdis of "organs", and that every such list may as well be replaced by the empty list. [M. Thompson 2008, p. 39]

It will prove illuminating to see what exactly the 'sub-metaphysical Scylla of "DNA" ' and the 'tautological Charybdis of "organs" ' consist in.

Thompson begins his considerations concerning a second-gear list of life-characteristics by discussing the possible list-occupant 'Living things are highly organized'.[322] Abstractly speaking, organization is not something one can measure on a scale of lower and higher degrees—'Is the administration of the University of Pittsburgh more highly organized than, say, a Buick or the Hope diamond, or more complex than the rules of chess?', Thompson rightly wonders [ibid, p. 36]. Abstractly speaking, then, the proposal at hand is empty. But perhaps sense can be made of it by supplying a relevant physical quantity—entropy, say. Surely, the configurations of matter constituting organisms exhibit lower entropy than other configurations of matter. Of course, as Thompson remarks, fresh corpses are thereby not excluded [ibid, p. 36]—and if they are, it becomes a mysterious question what entropic difference is required to distinguish between such corpses and their living counterparts. But even if, actually, all and only living things display such remarkably low entropic values, nothing is gained by using it to provide a determinate notion of organization for our candidate list-occupant. After all, though unlikely, it is not physically

[321] *Life*, n.d. In *Wikipedia*. Retrieved April 7, 2014, from http://en.wikipedia.org/wiki/Life.

[322] See M. Thompson [2008, ch. 2, §2]. Thompson takes his cue from a first-year students' biology textbook, which contains a list of the kind under consideration under the heading 'Signs of Life'—see Curtis [1979, pp. 20–1].

impossible that arbitrarily low entropy values are occasionally reached 'just so'. The problem, then, is that such a physical quantity *may* fail to coincide with the living, even if it *actually* does not. And it generalizes, of course, to proposals involving more than just the criterion of low entropy states. Extensional adequacy is simply not enough (otherwise a mere enumeration of all and only the living objects would suffice, we may add).

Interestingly, Thompson moves on to consider an even stronger suggestion. Suppose that we find out that life is only possible when DNA is involved—in other words, *no* combination of chemical substances that lacks DNA can constitute something that is alive. The requisite 'high organization' could then be cashed out very crisply in terms of containing DNA. Storrs McCall, for instance, writes:

> With DNA and protein-manufacture we have life: without them, merely physics and chemistry. [McCall 2012, p. 174][323]

Now, Thompson writes:

> The judgment about DNA, if it were true, would only show how resource-poor the physical world really is. It could make no contribution to the exposition of the concept of life . . . —except perhaps as pointing to a few gorillas and turnips might. [M. Thompson 2008, p. 37]

And indeed, looking more closely at what McCall says, we find that he has something different in mind than merely the presence of DNA and protein-manufacturing devices:

> [T]he division between living and non-living beings coincides with the introduction of informational software in the form of the genetic code. The hardware is the DNA and RNA molecules; the software is the encoded message they convey to the protein-making factories . . . . No coded informational software, no life. [McCall 2012, p. 175]

Surely, then, DNA plays a merely instrumental role in his definition of life. If no other kind of molecule was capable of embodying a code such that it could be used in the manufacturing of protein, then that indeed would merely show, as Thompson says, 'how resource-poor the physical world really is'. The notion of information involved is doing all the work.[324]

[323] McCall here follows Davies [1999].

[324] We should remark that both McCall and Davies in fact do not aim at a reductive understanding of life in life-less terms. Davies [2006] argues that there is 'downward causation' from the information encoded in the DNA to what happens physically, while McCall argues that, in addition, living things are governed by a special kind of analog information embodied in the structure of space-time—'[w]herever life exists, spacetime is filled with smaller, more detailed dynamic patterns that govern growth and development' McCall [2012, p. 181].

In the end, then, the problem is simply that DNA, or entropy, or any other purely physical quantity (or substance) is just that—something physical. There are other physical qualities or quantities as well: what makes DNA, or entropy, special beyond what characterizes them physically (or chemically)? The answer is plain: DNA, and low entropy states, play a special role *for living beings*—as, in the case of McCall and Davies, the role of carrying information. That is why we may reasonably use the presence of such physical quantities and stuffs to indicate the presence of life. Epistemically, then, they may be highly relevant, but not conceptually (and thus not metaphysically either). Hence, as Thompson says, such putative list-occupants are 'sub-metaphysical', they do not serve their intended purpose of contributing to an understanding of what life is.

On the other hand, something along the lines of organization, or of information for that matter, is of course central to life. The problem, however, is that spelling out what makes it central quickly leads us into a circle: organization or information is central to life only if it is understood as the type of organization or information that living things exhibit and use. Thompson writes with regard to such a conception of organization:

> [T]he relevant conception is simply equivalent to the idea of life: to be alive is to be *organ*-ized; to be alive is to be a subject of, say, 'vital organization'. [M. Thompson 2008, p. 38]

And this, then, is the 'tautological Charybdis' into which putative list-occupants must sink if they escape from the sub-metaphysical Scylla.

Surveying other putative list-occupants, we arrive at similar conclusions. Take, for instance, 'metabolism', and 'responding to stimuli'[325], which both occur in our Wikipedia-quote above. Thompson writes:

> Are we to say, for example, that the asphalt on a summer day "takes energy" from sunlight, and "converts" it into heat? And is an avalanche, on the other hand, the "response" of a snow-covered hillside to the "stimulus" of, say, excessive yodeling? [ibid, p. 39]

Again, the point is that the relevant conceptions of metabolism and responding to stimuli are either not making the relevant difference (which is the case Thompson highlights in this quote), or rather refer us back to life itself. The relevant conception of response to stimuli is, of course, response to stimuli *as it occurs in living beings*, that

[325] 'Response to stimuli' should here not be understood in a behavioristic vein, but rather broadly so as to encompass also phenomena of vegetative life—a plant's responding to incoming sunlight by manifesting the process of photosynthesis, say.

is, in the form of life-processes (as described in §4.4). And the relevant conception of metabolism equally involves the category of life-processes.[326]

Now, there is of course nothing wrong with either sort of statement about life. Living things do take and convert energy in the merely physical sense, just as Thompson's asphalt does. But that does not help someone who searches for a conception of life in life-independent terms. And living things are, of course, special in that their energy-conversion takes the form of life-processes. Again, that does not help someone who searches for a conception of life in merely second-gear terms: it is a conceptual truth from within the third conceptual gear.

To illustrate, recall our introduction of the third-gear concept of life-process in §4.4: life-processes always point beyond themselves to the wider context of the organism that hosts them. It makes sense to ask, given a partial description of a life-process, 'and what happens next'?, whereas such a question only makes sense for second-gear causal processes if it is specified which further influences are due. As Thompson writes:

> In a description of photosynthesis, for example, we read of one chemical process . . . followed by another, and then another. Having read along a bit with mounting enthusiasm, we can ask: "And what happens next?" If we are stuck with chemical and physical categories, the only answer will be: "Well, it depends on whether an H-bomb goes off, or the temperature plummets toward absolute zero, or it all falls into a vat of sulfuric acid . . . " That a certain enzyme will appear and split the latest chemical product into two is just one among many possibilities. Physics and chemistry, adequately developed, can tell you what happens in any of these circumstances——in *any* circumstance——but it seems that they cannot attach any sense to a question "What happens next?" *sans phrase*. [M. Thompson 2008, p. 41]

That there *is* an answer to such 'what happens next'-questions, then, illustrates precisely the sense in which life-processes point beyond themselves: they are processes that exist for a reason, a reason that goes beyond their merely physico-chemical components.

The search for a list of life, of the kind instanced by the quote from Wikipedia's 'Life' article, leads to the recognition that every single concept one might employ in its service—that is, in order to explicate what life is—turns out to already entail the very thing the list aims to capture: life itself. Organization, energy-conversion, having parts, responding to stimuli—all such concepts, *when used to characterize life*, form a kind of circle which one only enters by already having the concept of life. Thompson writes:

[326] Homeostasis, growth, and reproduction, three further list-occupants from our Wikipedia-quote, are easily seen to also depend on the notion of a life-process, or to be unspecific.

> A number of abstract categories——that of a concrete individual; of a thing's being a part of something; of order or organization; of one thing's following another in a process; of a thing's doing something——are all together determined or specified, or thrown into a higher gear, to yield the concepts: *organism*; *organ*, 'part' or 'member'; *vital order* or *organ*-ization; *life-process*; and *vital operation*. The abstract notion of existence . . . evidently bears the same relation to that of *life*: 'to be, for a living thing, is to live'. [T]his same shift of gear [turns] the abstract notion of a kind or of a 'natural kind' into that of a *life-form*——the notion, that is, of a living kind, or of a species (on one reading of that expression).
> These concepts, the vital categories, together form a sort of solid block, and we run into a kind of circle in attempting to elucidate any of them . . . . [M. Thompson 2008, p. 47]

We have provided our own sketch of this 'solid block' of life-involving concepts or 'vital categories' in §4.4, and we have called it, following Thompson, the third conceptual 'gear'.

In what follows, we examine the relation between third-gear life-processes and second-gear causal processes, in order to defend the reality of third-gear phenomena in a metaphysically robust way (which is, of course, the only possible way, given that we endorse the Aristotelian picture). Then, we move on to discuss attempts at reducing third-gear phenomena to second-gear ones, in order to show how and why they fail. And finally, we look at evolutionary theory and the challenges it appears to pose for our understanding of life-forms.

7.2 Mechanisms and Life-Processes

The second conceptual gear concerns persisting physical objects, their powers, and the processes they are involved in by manifesting those powers. On the basis of these conceptual materials, one can envisage situations in which such second-gear objects combine into more or less complicated systems capable of certain stable operations. Such systems can then be described as *mechanisms*. Interestingly, many life-processes, or at least those that take place on the biochemical level, are often described in mechanistic terms as well.

Our task in this section is to understand the relationship between mechanistic, second-gear thought and third-gear 'organic' thought. We do so by first looking at a bit of the history of mechanistic thinking (in §7.2.1 below), in order to understand its relationship to the Aristotelian hylomorphism that it replaced. Then we elaborate, in a variety of ways, on the relation and interaction between third-gear organisms and the second-gear materials from which they are made (in §7.2.2 below).

7.2.1 The Emergence of Mechanistic Thinking

The computer on which this essay is written illustrates a mechanism exhibiting remarkable complexity: successive pressings of keys initiate series of minute electrical processes resulting in patterns in the computer's memory representing the contents of this essay as it is being written, and also in patterns of light emitted from the screen representing the newly written part. Or, to give an easier example, the drivetrain system of a typical bicycle enables the transmission of power, exerted by the driver on the pedals, via the crankset over the chain to the (or a) rear sprocket, resulting in rotation of the rear wheel.

Such artifacts make clever use of the real possibilities to which the powers of the materials and objects of which they are made give rise. The mechanisms by which our artifacts function exploit these powers in specific ways geared towards our goals (e.g., writing essays, or transportation).

In the early modern period, fascination with the remarkable new mechanistic inventions of the time, such as the clockwork (which was developed in the 13^{th} century), encouraged a conception of (parts of) reality in purely mechanistic terms: the universe was likened to a big clockwork, a great mechanism governed by the universal and strict laws of physics, that was ticking its way into the future.[327] Inspiration for this picture was found, amongst others, in Newton's revolutionary discovery of just such universal laws of physics, of course. His results enabled strikingly accurate explanations and predictions of astronomical and tidal phenomena, and this made quite an impression, even though it took some time for his message to sink in, and even though his fundamental force of gravity, which was thought to work instantaneously over any distance, remained a mystery, not in the least for Newton himself. Still, Newton's work can also be seen as a culmination of a tendency that was already present for some time: a tendency to understand the universe mechanistically by use of sophisticated mathematical means. René Descartes, for instance, excepted only human beings from his mechanistic understanding of physical reality, and famously went so far as to construe animals as mere reflex-driven machines without a soul. And even in the case of humans, he understood the human body as just such a machine, which was special only in that it was partially controlled by the rational soul.[328]

[327] Earlier, we mentioned Laplace as one who thought along these lines—see §4.3, p. 147, and see also §6.2.6.

[328] See, e.g., the *Treatise on Man* [Descartes 1633, p. 108], the *Meditations on First Philosophy* [Descartes 1641, pp. 161–2, in the Replies to the fourth set of objections], and the *Passions of the Soul* [Descartes 1649, p. 329–30, §1.6]. (In the standard Adam and Tannery edition of Descartes's collected works [Descartes

These mechanistic tendencies have had a strong influence on the further development of scientific and philosophical thought over the centuries, culminating in broader acceptance of strongly mechanistic views on life towards the end of the 19th century. These views, and the mechanistic way of thinking on which they rested, were further nourished by scientific advances since then—in particular, by the 'modern synthesis' of evolutionary theory and Mendelian genetics, that was achieved in the early 1940s[329], and most famously, of course, by the discovery of the structure of the DNA molecule by Crick and Watson, documented in their [1953b,a], which marked the onset of molecular biology. The impressive insights that molecular biology has been able to achieve seem to constitute a compelling case for a mechanistic-reductive understanding of animate nature.[330]

This historical line of the emergence and consolidation of mathematical-mechanistic thinking, leading from early modern scientists and thinkers to fundamental questions in contemporary philosophy of biology, illustrates how second-gear thinking can become a metaphysical dogma that generates a Modern picture-based metaphysical realism, in this case a realism that takes reality-in-itself to be exclusively mechanistic (hence second-gear) in character.[331]

Now, there are those who are aware of what is happening behind the scenes here. Elizabeth Anscombe, for instance, complains about the way mechanistic thinking tends to determine our approach to understanding all of reality as follows:

> The high success of Newton's astronomy was in one way an intellectual disaster: it produced an illusion from which we tend still to suffer. This illusion was created by the circumstance that Newton's mechanics *had a good model in the solar system*. For this gave the impression that we had here an ideal of scientific explanation; whereas the truth was, it was mere obligingness on the part of the solar system,

1964–76], these passages can be found in, respectively, volume 11 p. 202, volume 7 pp. 229–30, and volume 11 pp. 330–1.)

[329] We will revisit the rise of this modern synthesis in §7.4.2 below.

[330] The discovery of Watson and Crick fairly quickly led to a sizeable literature on the question whether it enabled molecular biology to reduce classical genetics. Initially, the question was approached on the basis of classical models of theory reduction like that of E. Nagel [1961] and Hempel [1966]. Schaffner [1969] defended the thesis that the reduction was (going to be) successful, to which Hull [1974] replied with severe criticism (arguing, for instance, that the relation between classical genes and molecular genes is many/many), leading to a near consensus amongst philosophers of biology that reductionism of this kind is not plausible. See, e.g., Kitcher [1984], Rosenberg [1985, 1994], Dupré [1993], and B. Barnes and Dupré [2008]. (Though Rosenberg later departed from his critical stance towards theory reduction; see Rosenberg [2006].)

[331] Perhaps a disclaimer is in order here. We do not mean to provide a historically adequate explanation of contemporary habits of thought, nor do we pretend to be able to do so. Our interests are primarily systematic; even if the historical developments are much more complex and nuanced, it is the systematic relations between certain forms of thought, certain kinds of scientific research, and certain kinds of metaphysical convictions that we aim to highlight. Still, there are studies in the history of ideas that support the picture we here sketch in rough outline. See, e.g., F. Ellis [2005].

> by having had so peaceful a history in recorded time, to provide such a model. For suppose that some planet had at some time erupted with such violence that its shell was propelled rocket-like out of the solar system. Such an event would not have violated Newton's laws; on the contrary, it would have illustrated them. But also it would not have been calculable as the past and future motions of the planets are presently calculated on the assumption that they can be treated as the simple 'bodies' of his mechanics, with no relevant properties but mass, position and velocity and no forces mattering except gravity.
> Let us pretend that Newton's laws were still to be accepted without qualification .... We could say: of course nothing violates ... the laws of the force of gravity. But animals, for example, run about the world in all sorts of paths and no path is dictated for them by those laws, as it is for planets. ... [T]he laws are, rather, like the rules of chess; the play is seldom determined, though nobody breaks the rules. [Anscombe 1971, p. 143]

Anscombe's point is that, even if Newton's laws (or, we may add, any contemporary or future counterpart of it that may compete for the label of being 'completed physics') describes all there is to describe *concerning a given set of basic powers and properties*, such as gravity, this may very well be perfectly adequate without being the whole story about reality. Interestingly, she cites living beings (animals) as cases in point: it seems that she thought that the realm of the living differs from the physical realm in important ways—*without* thereby claiming that plants or animals do not conform to the laws of physics.

But how can it be that 'animals run about in all sorts of paths' independently of the second-gear laws, even though it appears that everything that happens inside them can be satisfactorily studied using the second-gear concepts of physics and chemistry? That is exactly the question we wish to assess in this section—attempts at reduction will be our target in the next section. Anscombe's analogy with the rules of chess provides an initial gloss on the issue. If the second-gear laws of nature are like the rules of chess, then the future is open in the sense circumscribed in §6.2.6 (and in §4.3, p. 140): at least some things have spontaneous or indecisive powers, and hence the future development of reality is determined only up to a certain range of real possibilities. That means that there is room for third-gear concepts—in particular, life-forms—to further narrow down that range of real possibilities, thus determining what will happen in ways that do not 'break the rules' of the second-gear powers.

At this point, one might think that we are heading in the direction of a certain kind of vitalistic understanding of life, which, if we read Dennett, seems not to be the most promising alley:

> [V]italism (the view that living things contain some special physical but ... mysterious stuff—*élan vital*) [has] been relegated to the trash heap of history, along

> with alchemy and astrology. Unless you are also prepared to declare that the world is flat and the sun is a fiery chariot pulled by winged horses—unless, in other words, your defiance of modern science is quite complete—you won't find any place to stand and fight for these obsolete ideas. [Dennett 1996, p. 24]

The thought is that if organisms influence what is going on in ways that go beyond the second-gear powers of their constitutive material parts, then surely there must be some hidden, ephemeral kind of force or stuff that produces this influence. And then the fact that empirical identification of this kind of force or stuff consistently fails puts the vitalist position in a quite unfavorable light. Vitalism is relegated to the 'trash heap of history' by noticing that mechanistic descriptions of life-processes seem, at least on the chemical level, possible throughout.[332]

We do not wish to enter into interpretive issues with regard to traditional vitalists here; rather, we wish to use the idea of vitalism *as understood by its enemies*, such as Dennett, to contrast it with our conception of life as a conceptually and hence metaphysically independent aspect of reality. Such a comparison helps to bring out what we have in mind. In order to do so, however, we need to get clear on where this understanding of vitalism comes from, and that, in turn, requires contrasting the early modern shift towards mathematical and mechanistic thinking we sketched above with the earlier Aristotelian paradigm. The scholastics taught that the things inhabiting reality were comprised of matter and form: the form organizes the underlying matter in accord with its own principle or essence.[333] In early modern thought, however, the strong tendency to embrace mechanistic thinking wherever that seemed feasible went hand in hand with abandonment of the very idea of such Aristotelian forms. We saw this happening in the case of Descartes: everything could be understood entirely in terms of matter alone, except for the influence of the human soul over its body.

Now, as we argued in chapter 2, this kind of thought rests on a mistake. Take, for instance, Descartes' ontologically fundamental notion of matter as defined by

[332] Vitalism was largely abandoned on such grounds after the first few decades of the twentieth century, but it was still a live option in those early decades—see, e.g., Driesch [1914]. We should point out, however, that a recently published collection of articles providing a balanced view on the history of vitalism over the past two centuries—Normandin and Wolfe [2013]—also contains some interesting considerations as to its contemporary significance that make clear that Dennett is massively overstating his case. See, in particular, Bechtel [2013] and Turner [2013]—the former, a dedicated 'mechanist', nevertheless urges that the vitalists were right in recognizing 'that the mechanist accounts [of their days] lacked the resources to account for some of the most fundamental features of living organisms' [p. 346], and tries to amend his mechanistic views accordingly; the latter, a dedicated 'anti-mechanist', urges that '[d]efining a new metaphysics of biology will mean engaging with and incorporating long-shunned "vitalist" concepts' [p. 287].

[333] And where the underlying matter itself involved forms, the distinction could be repeated, so that, at bottom, there was the infamous 'materia prima'.

extension, and qualified by motion, size, and shape: this conception of matter, if correct, would capture the true concept of matter, which is identical with its essence. And an essence is a form. A 'form-less matter' conception of reality is thus simply impossible: whatever we take reality to contain, we are bound to have corresponding essential truths, and these will constitute the relevant forms. So, the shift was not one from hylomorphism to materialism (in the sense of form-less matter). What was actually going on is, rather, that the traditional forms with which the Scholastics worked, which defined such things as chairs, human beings, frogs, and trees, were downgraded to mere inhabitants of reality as it is for us, while such highly general (and often questionable) concepts as Descartes' notion of matter were taken to characterize reality as it is in itself. That is, the attempt to leave hylomorphism behind was in fact never completely successful. It resurfaced in an almost unrecognizable way in the form of Modern picture-based metaphysical realism: only those concepts that made their way into the conception of reality as it is in itself were taken metaphysically seriously.[334]

Since the only kind of forms (i.e., concepts) that was thus taken to be still acceptable, after this early modern shift, was the kind that plays a role in mechanistic thinking, the result was that to understand reality only second-gear conceptual resources were allowed. Hence it is not surprising that those *rejecting* vitalism, because of their (explicit or implicit) adherence to this metaphysical restriction to mechanistic thought, could only frame that view in terms of something like a 'vital force' or 'vital stuff', understood somehow as analogous to familiar second-gear forces and stuffs such as gravity and water.[335] (To repeat, whether or not this was what the *defenders* of vitalism had in mind is a question we will not touch upon here.)

It will be clear, by now, that it is our aim in this section to get rid of the confusion that was installed when the Modern picture took over, and revisit the hylomorphism of the Aristotelian tradition (which we have been defending on a very general level in §2.4). By doing so, we not only arrive at a reasonable, non-mechanistic understanding of third-gear phenomena, but also at a better understanding of second-gear phenomena.

[334] F. Ellis [2005, p. 1] calls the confusion over the role concepts play in the understanding of reality in itself, which made it seem as if hylomorphism was really rejected, 'the syndrome'. As we discussed in §5.2.4, the Humean extreme to which this tendency led really does reject hylomorphism because it rejects the very idea that our concepts capture aspects of reality itself—that is, on this extreme we end up with either skepticism or metaphysical anti-realism (see §1.2).

[335] In those early twentieth century decades, there were others who attempted to steer a middle way between mechanistic reductionism and vitalism by claiming that while no extraneous forces were required, the unity and organization of the individual organism should be taken as basic. See, e.g., Woodger [1929] and Haldane [1931].

7.2.2 Matter and Material

We claimed, in §4.3, that second-gear objects fall under fundamental kinds that are generically characterized by their being subject to the second-gear forms of thought. Suppose that electrons form such a kind: then, we may say that this particular electron *can move* as it does *because* that accords with the conceptual truths attached to the concept *electron*—for instance, in accord with the conceptual truth that electrons are negatively charged, where such charge is understood as one of its characteristic powers. As we observed, the 'because' here is not a causal one in the usual sense: we adopted an Aristotelian name for it, *formal causation* (see §4.3, p. 139, and also §6.1.4, p. 199). It just means that this electron behaves in accord with its nature. By contrast, we may also say that the electron *moves* as it does because it has been shot into a cloud chamber in which a certain electromagnetic field is present. The 'because', in this case, *is* a causal one in the usual sense—in Aristotelian terms, it is an efficient cause. The efficient cause provides the occasion for the electron to express aspects of its nature by way of formal causation.[336]

Once we realize that mechanistic thinking requires such a notion of formal causation—why do the various components of any mechanism act and interact as they do?—we have made one crucial step towards understanding the way in which the realm of the living may influence what is going on without requiring vital forces or stuffs of the type Dennett has in mind when he rejects vitalism. For the third-gear forms under which living organisms fall, as sketched in §4.4, operate just as the second-gear forms, such as that of our electron, do: by way of formal causation. Why is this tree growing leaves of just that particular five-fingered shape? Because it is a sugar maple, and growing leaves of that shape accords with what it is to be a sugar maple, that is, with the conceptual truths attached to the concept *sugar maple*. The 'because' again indicates formal causation. And, again, we may distinguish this sense of cause from the efficient cause, which in this case is the change of season, with its increase in temperature and daylight.[337]

Of course, this observation immediately leads to the complementary observation that formal causation is also much different in the case of third-gear life-forms than it is in the case of second-gear natural kinds. Third-gear life-forms express themselves

[336] To repeat, we do not intend to be faithfully representing Aristotle's doctrine of causes; our interests are systematic, not exegetical, and our use of Aristotelian labels is thus meant as expressing allegiance to a broader philosophical orientation, not to a specific metaphysical doctrine. See our brief remarks in §4.3, p. 139.

[337] We refer those worried by the threat of emptiness in explanations of this rough form (why is this X doing A? Because that is what Xs do) to our brief remarks concerning that worry in §6.1.4, p. 199.

in the organization of underlying second-gear *matter*, which is thus turned into *material* for the organism in question. Indeed, that is one crucial part of any living creature: incorporating material from its surroundings by way of life-processes that are typical of its life-form (see, for instance, our brief sketch of part of the life-process of photosynthesis in §4.4). No such relationship between object and underlying matter is present on the second conceptual gear—our electron does not assimilate underlying materials in order to sustain its existence.

Notice that this understanding of the relation between living organisms and the second-gear stuff they are composed of falls under the heading of 'emergentism', at least under one reading of that notion, provided by Broad:

> Put in abstract terms the emergent theory asserts that there are certain wholes, composed (say) of constituents A, B, and C in relation R to each other; that all wholes composed of constituents of the same kind as A, B, and C in relations of the same kind as R have certain characteristic properties; that A, B, and C are capable of occurring in other kinds of complex where the relation is not the same kind as R; and that the characteristic properties of the whole R(A,B,C) cannot, even in theory, be deduced from the most complete knowledge of the properties of A, B, and C in isolation or in other wholes which are not of the form R(A,B,C). [Broad 1925, p. 61]

If one reads 'relation (of the same kind as) R' to mean 'life-form', 'constituents A, B and C' to mean 'second-gear stuff', and 'certain characteristic properties' to mean 'third-gear features and life-processes', one gets more or less what we have proposed. For then the 'most complete knowledge of the properties of A, B, and C in isolation' will amount to second-gear knowledge, and from such knowledge it is indeed impossible, 'even in theory', to derive third-gear conclusions.

Bedau [1997, 2002] draws a distinction between 'weak' and 'strong' emergence. The former amounts to the impossibility of deriving the weakly emergent features from the properties of the components in conjunction with the (reductionist) thesis that, even so, the behavior of the parts fully determines these weakly emergent features. There is, then, nothing metaphysically special about such weakly emergent features; weak emergence is, at bottom, an epistemic notion. By contrast, strong emergence is weak emergence without the reductionist thesis. The emergentism we are in effect defending is, then, of the strong variety.

Now, in the words of Powell and Dupré [2009, p. 59], this implies 'unpredictability in principle, ... obstacles to prediction somehow inherent in the nature of things', which is 'often seen as philosophically mysterious'. Given the distinction between the second and the third conceptual gears, the supposed 'mystery' is, however, readily explained: without a shift of gears, you won't get from the second to the third

conceptual gear. And the emergence we are defending consists precisely in that shift in gears.

Where strong emergence is in the offing, 'downward causation' is bound to be nearby. So, for instance, in the case of Dupré, who argues against reductionism and in favor of strong emergence [see, e.g., Dupré 2010, §3]. He writes:

> Downward causation seems a very natural way to think of much of what I have been saying about molecular biology. What causes the human genome to behave in the particular ways it does—for example, various sequences being transcribed or not at varying rate, changes in conformation and spatial relation of chromosomes, and so on—is a variety of features dispersed over the surrounding parts of the cell. The behavior of the part is to be explained by appeal to features of the whole. . . .
> The cell, I think we must say, with all its intricate structure and diverse contents, is what causes these contents to behave in these life-sustaining ways. [Dupré 2010, p. 42–3]

For Dupré, then, downward causation is causation that flows from the whole to the parts: the cell causes its components to behave the way they do. We are sympathetic to this thought, but find it to be confused: how does such downward causation by 'the cell' work? Our proposal, of course, is to make sense of the idea in terms of *formal causation* (see p. 270 above): the cell organizes its underlying material in accord with *what it is*—in the case of a unicellular organism, that will be the relevant life-form, in other cases, it will be the relevant aspect of the life-form which the encompassing multicellular organism instantiates.[338]

Our proposed understanding of the relation between second- and third-gear phenomena is thus thoroughly anti-reductionist, emergentist in spirit. As such, it is bound to invite difficult (skeptical) questions, of course, and we can learn a few interesting things about the implications of our proposed understanding by reflecting on some of them. We discuss (1) a skeptical thought of a mereological variety, (2) a difficulty concerning 'life-like' chemical processes, (3) a further worry concerning 'as-if' teleology, and, finally, (4) a systematic worry to the effect that third-gear formal causation may well be incompatible with second-gear indeterminism, contrary to what we claim.

[338] Dupré at times seems to come close to the kind of view we are proposing, yet he never really distinguishes between second- and third-gear thought. The result is that he can only express his anti-reductionism in mereological terms—that is, in terms of 'the whole' influencing 'the parts'. He frequently alludes to the importance of 'the wider context' for understanding biological phenomena, in particular on the molecular level, and is positively impressed by 'systems biology'. See, e.g., Dupré [1993, 2010, 2012, 2013].

1. A Mereological Difficulty Second-gear objects can be composite, and our example of the electron is therefore potentially misleading. What about, say, a helium-4 atom? Are not the composite pairs of neutrons, protons and electrons the material for that atom in much the sense in which organisms are built up from materials, just on a much smaller scale? Now, it is far from clear how exactly the individuation of such subatomic particles works, and hence what sense can be made of the very idea of the parts of atoms.[339] It may be that our helium atom is indeed a system comprised of six separate subatomic particles, as pictures of such atoms in popular scientific magazines (and high school textbooks) suggest, in which case the whole is a result of the interactions of those particles. It may also be that the helium atom was, e.g., generated out of those particles, and can be decomposed into them, but is itself somehow a unity of a different sort, obeying its own laws. And similar options can be conceived of when it comes to molecules, and also when considering larger collections of atoms or molecules, mixtures thereof, etc.[340] But we should settle such questions in conversation with the physicists, of course, not from the armchair. The crucial thing to note, here, is that all of the above situations can be understood entirely in second-gear terms: we either have systems of interacting second-gear objects (which we may take together under a single non-fundamental concept) or single second-gear objects falling under some fundamental concept—a natural kind. The properties of such single second-gear objects may still be understood in part in terms of their components.[341] The laws of various composite kinds of things take their place within the overall system of second-gear kinds, their powers, and the corresponding laws. They fall under the second conceptual gear, that is, not under the third conceptual gear—and that is what distinguishes them from the living.

[339] And even the atoms themselves seem to create difficulties with respect to their individuation, due to their double standard of existing both as particles and as waves. Consider the following illustration, provided by French [2011, §2]. In classical statistical mechanics, given two intrinsically indistinguishable individual particles a and b, and two positions P_1 and P_2 they can both occupy (in an abstract sense of 'position'), there are four possible situations: (1) both a and b occupy P_1, (2) both occupy P_2, (3) a is in P_1 and b is in P_2, and (4) b is in P_2 and a in P_1. (3) and (4) are separate possibilities only because a and b are distinct *individuals*. Yet in quantum statistics, things are different. For systems with symmetric state functions, to which Bose-Einstein statistics applies (such as systems of helium-4 atoms, which are bosons) there are only three possible situations (both in P_1, both in P_2, or one in P_1 and one in P_2)—there is literally no difference between (3) and (4). And for systems with an anti-symmetric state function, to which Fermi-Dirac statistics applies, there is only one possible situation (one in P_1, one in P_2). As some have argued, these facts imply that such particles cannot be individuals—see, e.g, Post [1963]. On the other hand, others have argued that such particles can be individuated after all, albeit in unusual, weak ways (e.g., relationally)—see, e.g., French and Redhead [1988], Muller and Saunders [2008], and Muller and Seevinck [2009]. For a good overview of this issue of 'quantum individuality', see French and Krause [2006] and French [2011].

[340] The question is, ultimately, what the fundamental, conceptually basic categories are—see chapter 2.

[341] As we remarked in §4.3, p. 138, we ignore stuffs, to prevent needless complication.

Reflecting on these thoughts a little more, one can come to appreciate the phrase 'periodic *system* of elements': a given chemical element is not only essentially characterized, as defenders of 'a posteriori metaphysical necessities' are wont to point out, by its atomic number and everything that follows from that with regard to the element itself, but also by its standing in systematic relations to the other elements, its diverse possibilities of interaction with them, etc.[342] Second-gear thought aims at an ever more encompassing understanding of second-gear objects, such that eventually it embraces the whole cosmos—see, again, the Laplacean fantasy briefly discussed in §4.3, p. 147. Mereological relations of various sorts are simply part of that overarching second-gear understanding.

2. Life-like Chemical Processes Now for the second issue, that of life-like chemical processes: these are processes that resemble life-processes in certain ways but occur outside of the realm of the living. If the third conceptual gear were to apply also to such processes, it would seem that we failed to capture what is distinctive of life by providing our sketch of the third conceptual gear—in which case the skeptical thought that there is nothing 'distinctive' out there to capture looms large.

Take, for instance, a candle flame. It takes up underlying material (say, paraffin wax) by melting it, which allows it to travel up the wick by capillary action, where it evaporates and then burns, sustaining the flame itself. The flame also takes in oxygen in its (multistage) process of burning, and releases carbondioxyde and water vapor (and small quantities of some other products). The flame, thus, seems to be a unity that takes up material from the environment, which it transforms and consumes to sustain its own existence.[343] So why is the flame not a living organism, in our sense? The answer is plain: the entire process in which the flame consists is a second-gear process that can be completely understood in terms of the powers of the objects and stuffs involved.

The obvious reaction to this answer is, of course, the following: isn't the difference between such a flame-process and, say, the complex of life-processes of a sugar maple then not a mere difference in complexity? If so, organisms are nothing but very complex flames, so to speak. Now, this thought illustrates the second-gear metaphysical

[342] Recall that 'a posteriori metaphysical necessities' simply are conceptual truths, on our construal of those. Their 'a posterioricity' is not of much importance—see §3.3.3.

[343] We ignore the fact that candles are artifacts, and thus involve the purposes of human beings—artifacts are involved not just in the second conceptual gear but in higher gears as well. (We ignored the artifactual status of balls for the same reason in §4.3, p. 137.) If you don't like the example, you could substitute Mount Wingen for it—the famous 'Burning Mountain' in New South Wales, Australia, where an underground coal seam fire has been going on for thousands of years.

dogma: we have decided in advance that reality only contains second-gear objects and processes, and hence our view of life should conform to that restriction.[344] Yet that does not at all conform to our actual science of living things, i.e., biology—despite the strong tendency within that discipline to look at molecular biology as the hopeful bridge towards a fully reductive view that vindicates the dogma.

It is true, on our view, that life-processes can be biochemically described. We repeatedly cited the example of photosynthesis as a case in point. Another example, frequently discussed in the contemporary literature over the question of biological reductionism, is protein synthesis, which we now briefly describe to bring out a few interesting oddities that are absent in cases like our candle flame. Part of the process of protein synthesis is *protein folding*, which proceeds by a complex, but relatively well-understood mechanism (though Dupré describes the issue as '[a] major problem in molecular biology' [Dupré 2010, p. 42]). The so-called translation process, guided by a ribosome (a dedicated organelle), produces a sequence of amino acids on the basis of a strand of mRNA which was transcribed from the DNA in the cell nucleus. The protein thus formed assumes the three-dimensional structure to which its particular chemical structure, together with the environmental conditions (such as pH, temperature, concentration of salts) gives rise. That process is often catalyzed by special proteins called chaperones. It may undergo further changes in order to attain the three-dimensional form on which its functionality depends. These changes include: forming specific internal bonds (disulfide bridges), cutting the chain at some point or other, removing certain amino acids from the end of the formed chain, fusing together various already folded proteins into a whole, and forming metal clusters.[345] At all stages, the process of protein synthesis involves dedicated organelles, proteins, and other components—which is to say that there are reasons for them to be around during the process. The same cannot be said in the case of a flame that consumes some supporting material—the oxygen and the wax are not around *in order for* the burning process to take place.[346] Moreover, these organelles, proteins, and other components are themselves products of complicated processes of biogenesis, involving further organelles and components (and, of course, the process of protein synthesis itself).[347] The life-process of protein folding is thus embedded

[344] Recall our discussion of 'organization' as a defining characteristic of life in §7.1 above.

[345] See Whitford [2005, esp. chs. 8 and 11]. Hüttemann [2011, §4.3] contains a good discussion of this example from a philosophical perspective. Relevant references to empirical studies of the process can be found there as well.

[346] Recall that we are ignoring the fact that candles are artifacts.

[347] A notorious case in point is the production of the relevant mRNA-strand, which codes for the protein that is to be produced—since genes are often distributed over various parts of the total DNA, the tran-

in a wider context of life-processes to which it is intimately connected. Again, that cannot be said in the case of the flame. This point becomes especially vivid once we broaden our view and contemplate the fact that, as the case may be, the protein folding process we started with occurs because that specific protein is necessary for an immune response to a virus that has invaded the tissue within which the hosting cell is located. Again, nothing of this sort is visible in the case of a flame.

3. 'As-If' Teleology It is, of course, true that one may use third-gear terminology to describe a flame. On closer inspection, however, we see that that is a mere projection on our part: the flame is not taking in flammable material 'in order to' sustain its further existence—the material is simply burning, and that's all. The same cannot be said in the case of life-processes such as protein folding or photosynthesis. The case is, then, much like our habit of psychologizing our own technology: 'The printer *refuses* to print my text!'. Nothing of interest follows, from such examples, for the relation between second- and third-gear phenomena.[348] Again, to believe that our application of third-gear thought is *always* a projection, in the case of the flame as well as in the case of 'real life', is to endorse a groundless second-gear metaphysical dogma—groundless, that is, because it rests on a form of skepticism concerning something that is very much open to view: life.

Similarly, we can be very short on thoughts like the following: might it not be that everywhere we perceive the realm of the living *as if* it were characterized by such an autonomous style of organization, while in reality there is nothing but second-gear processes? We ask: what reason do we have to think so? Only a preformed conviction that reality really only contains second-gear processes, not third-gear ones, can inspire this skeptical thought of cosmic illusion.[349]

scribed RNA has to be 'spliced' at the proper places to remove the non-coding parts ('introns'), an activity carried out by what is called the spliceosome (Rogozin *et al.* [2012] contains interesting information on this peculiar entity). Another notorious case in point is the production of organelles. For instance, it is unclear for several kinds of organelles whether these can be produced from scratch in a cell, or rather require existing organelles to be copied from. See, e.g., Mullins [2005]. A rich source of further examples, more or less related to the case at hand, can be found in the works of Dupré—see fn. 338 on p. 272 above for references.

[348] It should be no surprise that we thus disagree with Dennett's views on these matters—as voiced in, e.g., Dennett [1987]. We find Dennett's views on 'design' and 'intentionality' to be as close to projectivism as Sellars's understanding of what he calls the 'manifest image'—see §1.2, p. 34.

[349] Such a conviction has been expressed quite explicitly by, e.g., Smart [1959], who writes that 'biology is related to physics and chemistry in the way in which radio-engineering is related to the theory of electricity and magnetism, and so on' [p. 363]. On the other hand, the idea of 'as if'-teleology has famously been defended, within a rather different context but in part on strikingly similar grounds, by Kant [1790, see esp. §75]. Kreines [2008, p. 345] provides a nice summary of Kant's view: 'Kant actually argues that we necessarily conceive of living beings in irreducibly teleological terms. But we cannot know that living beings themselves truly satisfy the implications of teleological judgment. . . . And this skepticism requires

4. Third-Gear Formal Causation and Indeterminism We turn, now, to the final worry, of a systematic nature, which needs to be laid to rest. Although the second-gear rules are not broken when third-gear life-forms exert their formal causal influence on what happens, still, so the worry goes, such influence by the organism on what is happening somehow breaks statistical laws telling us that, within the range of real possibilities, what happens is random.

That third-gear life-processes should somehow break the randomness that is dictated by the second-gear level is a thought that Anscombe also briefly discussed in her inaugural lecture, albeit with regard to the human will rather than with regard to life. She writes:

> Certainly if we have a statistical law, but undetermined individual events, and then enough of these are supposed to be pushed by will in one direction to falsify the statistical law, we have again a supposition that puts will into conflict with natural laws. But it is not at all clear that the same train of minute physical events should have to be the regular correlate of the same action; in fact, that suggestion looks immensely implausible. It is, however, required by the objection.
> Let me construct an analogy to illustrate this point. Suppose that we have a large glass box full of millions of extremely minute coloured particles, and the box is constantly shaken. Study of the box and particles leads to statistical laws, including laws for the random generation of small unit patches of uniform colour. Now the box is remarkable for also presenting the following phenomenon: the word "Coca-Cola" formed like a mosaic, can always be read when one looks at one of the sides. It is not always the same shape in the formation of its letters, not always the same size or in the same position, it varies in its colours; but there it always is. It is not at all clear that those statistical laws concerning the random motion of the particles and their formation of small unit patches of colour would have to be supposed violated by the operation of a cause for this phenomenon which did not derive it from the statistical laws. [Anscombe 1971, p. 146]

Anscombe's analogy is rather enigmatic, but the underlying thought is clear enough. Applying it to our case of living beings, we get the following: we have the laws that the particles in the glass box obey—these stand for the underlying second-gear components. Then we have the fact that these particles are constantly being arranged in such a way that the word "Coca-Cola" is formed out of patches of similarly colored such particles—that fact stands for the fact that the underlying second-gear components happen to be doing just what is necessary for the organism to manifest and maintain itself. And just as the "Coca-Cola"-phenomenon appears in a large variety of ways, without breaking any law of the particles—neither deterministic

Kant to carefully limit his positive claims about teleology: it is subjectively necessary we conceive of living beings in teleological terms, and this conception is legitimate when employed not as an explanation but as a heuristic aid for scientific inquiry.'.

ones nor statistical ones—so the living organism is dynamically present without breaking any of the second-gear laws of its second-gear components.

In the end, the point is this: we notice a plethora of third-gear phenomena around us—phenomena that, like the "Coca-Cola"-phenomenon of Anscombe's analogy, make use of underlying material for purposes that are alien to that material taken separately (as the phenomenon of decay upon death shows). However, unlike the "Coca-Cola"-phenomenon, which does not really make sense, we can actually understand how and why organisms do what they do. The whole teleological order of explanation that is so characteristic of living things comprises an aspect of reality—its third-gear aspect—that is autonomous with respect to its second-gear aspect, yet is intimately connected to it, as it only becomes manifest on the basis of underlying second-gear stuff.[350] The second-gear components determine a range of second-gear real possibilities, which are completely independent from the wider context of the life-form in which they are involved. A given amino acid molecule can, of course, bind to another amino acid—but that is just one of its many chemical possibilities. This leaves room for the life-form to exploit those possibilities, thereby organizing the materials at its disposal in accordance with its nature. Recall Anscombe's chess analogy: the second-gear determination of what is really possible is like the rules of chess. Living organisms make chess moves in accord with these rules.

7.3 To Reduce Or Not To Reduce

We have already stressed several times what our basic reply to attempts at reducing third-gear phenomena to second-gear phenomena consists in: the motivation for such attempts can only come from a pre-installed conviction to the effect that only second-gear thought is metaphysically acceptable. Insofar, such attempts betray allegiance to Modern picture-based metaphysical realism. Given the obvious fact that we understand the living by use of typical forms of thought and explanation that are clearly different from what we find in physics and chemistry, such a metaphysical dogma can only be maintained by incorporating a reductive (or eliminative) approach to animate nature.

[350] Compare, e.g., computers: it is highly unlikely, speaking purely physico-chemically, that our world's stuff should be arranged as it is in our computers so frequently. From a second-gear point of view, nothing further can be said about it—the coming about of these computers in no way offended the laws of nature, of course. But we *can* understand why they are there if we understand them as artifacts made by humans for specific reasons—that is, if we stop insisting on using only second-gear thought and allow for higher-gear thought as well.

More interesting than merely insisting on such a highly general defense of third-gear thought, however, is to consider particular attempts at reduction, in order to see where and why they go wrong, by our lights. That is the aim of the present section. We choose, by way of case study, the work of Ruth Millikan. Given that Millikan aims to contribute mainly to philosophical disciplines that fall outside the scope of the present essay (such as philosophy of mind and language), it may be surprising to focus on her work in this context. However, it is precisely because she is at least as much interested in the ramifications of third-gear forms of thought for a theory of mental and linguistic content or meaning, as she is in a respectably reductive understanding of those forms of thought, that we discuss her work here—she clearly does not intend to downplay third-gear forms of thought.

We need, by way of stage setting, to understand the basics of Millikan's overall views. That is what §7.3.1 aims at. Thus armed, we examine Millikan's metaphysics in §7.3.2, in order then to move on to a critical assessment of her reductive understanding of third-gear concepts in §7.3.3.

7.3.1 Millikan on Abilities, Unicepts, and Representations

In her afterword to the recent *Millikan and Her Critics* volume [2013], Millikan writes:

> What I have been talking about all these years . . . are not really empirical concepts at all, but *unicepts*—"uni" for one, of course, and "cept" for Latin *capere*, to take or to hold. A unicept takes in information gathered from many proximal stimulations and holds them as one distal object, property, kind, and so forth. Unicepts *should not be confused with concepts* and I propose to do so no longer. [Millikan 2013, p. 281]

These remarks are helpful, for us, because they prevent confusion of what we call 'concepts' with what Millikan used to call 'concepts' up until very recently. We therefore adopt Millikan's suggestion for our discussion of her views, and replace 'concept' with 'unicept' even where it occurs in quotes from her own earlier works.[351]

As the quote already suggests, unicepts are indeed quite different from our concepts: unicepts are, in part, abilities to reidentify the same in different circumstances. They are part of the 'mental machinery' of individual organisms, and hence not the essences of what is thereby grasped (though this claim will turn out to require qualification; see §7.3.2 below). Thus, unicepts are loosely related to what we have called 'conceptions' (graspings of concepts; see §1.3, p. 44). They are part of the *activity* of thinking, not part of its *content*. Millikan writes:

[351] For more on the shift from 'concept' to 'unicept', see Millikan [2012, esp. pp. 228–9] and Millikan [2014, esp. pp. 14–7].

> A unicept is a specific individual faculty developed for a very specific purpose, the purpose of collecting and integrating information about some particular thing. . . . [U]nicepts are not things that people share. Each of us has our own private stock of unicepts. Many of your and my unicepts do, of course, succeed in gathering up information about exactly the same things in the world, but they do this, pretty unexceptionally, in somewhat different ways, often utilizing many overlapping input methods but also many that are distinct. (Helen Keller's unicepts succeeded in gathering information about many of the same things yours do, but in ways most of which were very distinct from your ways.) [Millikan 2014, p. 16]

Of course, Millikan does have a lot to say about content, too. In particular, she defends an externalist account of unicept meaning: unicepts refer precisely to those real-world things or kinds (Millikan calls them 'substances' in one presentation of her views—see below) which it is their task to reidentify, and for the rest they do not have any canonical content—the information gathered by one person, with the help of a given unicept, may very well be entirely disjoint from the information gathered about the same thing by some other person. It is the *gathering* of information, not specific information *gathered*, that forms the core business of unicepts (although, again, qualifications will have to be made later on).[352] Millikan writes:

> [T]he task of [unicepts] is to enable us to reidentify substances through diverse media and under diverse conditions, and to enable us over time to accumulate practical skills and theoretical knowledge about these substances and to use what we have learned. [Millikan 2000, p. 2]

By using the term 'substance', Millikan intends to talk not only of particular objects but of anything one could have a unicept of in this sense. Substances are thus defined primarily epistemically:

> Substances are those things about which you can learn from one encounter something of what to expect on other encounters, where this is no accident but the result of a real connection. [ibid, p. 15]

In other words, substances are such that induction works: you can learn from one encounter what to expect on the next because you are encountering *the same substance*. This kind of induction applies to a wide variety of cases, making Millikan's category of 'substance' very broad. Unicepts not only hook onto individual things (Aristotle's primary substances), but also onto what Millikan calls 'real kinds' (Aristotle's secondary substances), where these include both 'eternal kinds', for example

[352] The relevant externalist tradition started, of course, with Kripke [1972] and Putnam [1975]. Millikan, however, argues that both Kripke and Putnam did not really manage to put forth a truly externalist account of meaning—see Millikan [2000, §3.5].

chemical kinds, and 'historical kinds', such as biological species.[353] In short, then, substances are entities that *have* properties one could learn. We turn to the specifics of the underlying metaphysical picture below (§7.3.2).

We now first look at the question what, then, an ability is, according to Millikan. An obvious proposal comes to mind: abilities form a subclass of dispositions or powers—that is, they are at home within the second conceptual gear. However, Millikan contrasts abilities with mere dispositions as follows:

> Abilities are distinct from dispositions in having a necessary involvement in the purposive and nonaccidental order. [Millikan 2000, p. 58]
>
> [W]hat I have an ability to do is what my systems were maintained or selected for doing. ... To an ability there always corresponds a disposition, but it does not follow that an ability IS a disposition. ... There is no such thing as *the* set of conditions under which it is necessary that any person be able to manifest their ability to do word processing or to swim. [ibid, p. 61–2]

The 'purposive and nonaccidental order' is, of course, the third conceptual gear. In effect, Millikan is doing with 'ability' precisely what Thompson does with concepts like 'organization' and 'metabolism' (see §7.1 above): the merely second-gear variant of the notion ('disposition') naturally has something to do with abilities, but is simply placed on the wrong level when one is trying to understand abilities—just as entropy and energy-conversion turn out to be simply wrongly placed when one is trying to understand life. The difference, however, is that Millikan does hint at a reductive understanding of her 'nonaccidental order': she equates it with the developmental and evolutionary history of the 'systems' involved. We critically consider that reductive understanding in §7.3.3 below.

You have an ability to swim in virtue of a disposition defined through your individual past. That means that more persons can have the *same* ability (as individuated by the target activity) resting on radically *different* dispositions—which is just the structure we also found in our sketch of Millikan's unicepts above. What is crucial is the teleological aspect of abilities: they have *functions*, they result from systems that have been *designed* to yield univocal unicepts—designed, that is, either by natural selection or by learning systems. That is the 'purposive and nonaccidental order'. Millikan's view thus requires elimination of the anti-third-gear conviction that biological items only have 'purposes' in a metaphorical sense at best.[354] As we said, that is precisely what makes Millikan's account interesting for our purposes: she really

[353] For Aristotle's terminology, see his *Categories* [1963, pp. 5–6; 2a11–18]. See, e.g., Millikan [2000, chs. 1–3] and in particular Millikan [2010] for more details on why unicepts for individuals and for kinds can be treated similarly, from the point of view of cognition.

[354] Millikan argues against this intuition at length in the first two chapters of her [2004].

wants to take third-gear concepts seriously, puts them to good use in her overall philosophy, yet also provides a reductive take on such third-gear concepts.

Millikan's account of what abilities are lies at the very core of her overall position: it determines to a large extent what unicepts are. Your unicept of, say, water, is, in part, an ability to *reidentify* water, that is, to notice that certain information coming in through your senses (or through language) concerns this same stuff again.[355]

The significance of abilities to reidentify, in the sense Millikan develops, is biologically obvious: they play an important role in directing behavior purposefully. It is quite useful for animals in general to recognize the same again under different circumstances; i.e., to recognize dangers, food, possible mating partners, etc.

As an aside, notice that what we witness here may well be a new kind of metaphysical dogma at work. Suppose, for a moment, that thought, language, intentional action and the like comprise a higher conceptual gear (a fourth, or perhaps a fifth, depending on whether sentient life requires a gear of its own). Then, Millikan can be taken to endorse the view that these higher-gear concepts can be understood satisfactorily (and reductively) in third-gear terms: thought, language and the like are just instances of biological phenomena exhibiting a remarkable degree of sophistication and organization (compare the idea of a 'list of life' that Thompson highlighted—see §7.1). In fact, the title of Millikan's first book, *Language, Thought, and Other Biological Categories* [1984], quite explicitly expresses this reductive stance.[356] The questions whether there really are such higher conceptual gears, and if so, what exactly we should then think about Millikan's position on this point, lie beyond the scope of the present essay.

We have now seen enough of Millikan's overall framework to get an idea of what is going on, in particular of what role third-gear concepts play in her approach to cognition, language, perception, and meaning. Her approach to such issues in the philosophy of mind and language, through the use of third-gear thought, makes the resulting view much more plausible and successful than approaches that rely on purely second-gear thought—which is not surprising, on our view, given that the phenomena she studies are at least third-gear phenomena, if not higher-gear ones, and hence are easier to understand from a third-gear than from a second-gear

[355] Indeed, Millikan argues that information coming in through language is in many respects the same as information coming in through the senses: understanding talk about something is a form of perception of that thing, on her view. See Millikan [1984, ch. 9, 2000, ch. 6, 2004, ch. 9, 2005, ch. 3].

[356] Elsewhere, Millikan writes: '[A]ll of the basic norms applying to cognition are biological norms' [Millikan 1993, p. 3]. As we said, Thompson develops a radically different understanding of intentional action as forming part of a higher conceptual gear, distinct from the third gear. See M. Thompson [2008, esp. Part Two].

perspective.[357] Let us therefore move on to consider the metaphysical foundations on which the view is based.

7.3.2 Millikan's Metaphysics

As we saw, Millikan approaches what she calls substances through unicepts, that is, she takes substances to be entities that can be reliably reidentified. Unicepts are abilities to reidentify, which is why Millikan's account can be classified as an externalist, direct reference view. But, interestingly, she adds a further ingredient to her account of unicepts, which we have omitted up to this point. That further ingredient consists in what Millikan calls 'substance templates':

> To have the unicept of any individual person, you must know what kinds of questions can be asked ... about individual people .... The primary interest of groupings like *persons*, *species*, and *chemical elements* is ... that they bring with them "substance templates". ... *Physical object* seems to be a pure substance template. ... With rare exceptions, however, categories that bring with them substance templates also bring at least a bit more. They correspond to substances displaying at least a few common properties as well as bringing substance templates with them. [Millikan 2000, p. 10]

This is Millikan's version of what we have argued for in chapter 2, but from an epistemic perspective: if you know *that* something is, you have to know *what* it is—or, at least, have a minimal clue as to what it is. You cannot (re)identify something if you do not know any substance template under which it falls.[358] Now, Millikan is no adherent of the Aristotelian picture—in fact, as we will see shortly, her view aptly illustrates Modern picture-based metaphysical realism. Therefore, there is no straightforward move from the role substance templates play for our cognitive economy (i.e., for our unicepts) to their ontological underpinnings—no identification

[357] One interesting illustration of this claim, which however does not seem to have been noticed all that much, is Millikan's solution to the so-called 'Kripke-Wittgenstein paradox', which was introduced by Kripke in his [1982] (Sprevak [2008] does mention Millikan's proposal in his discussion of that paradox, but misunderstands her position entirely.). The paradox is a skeptical challenge: it states that no determinate meanings can be ascribed to our thoughts and words, because they have only been applied (and will only be applied) in a finite number of cases, which fail to settle any determinate rule for correct application. As Goldfarb [1985, esp. p. 476] in effect observes, Kripke relies in the construction of his paradox purely on second-gear thought—e.g., he observes that no *disposition* can do the trick, since dispositions are purely factual and lack the required normative force (see also Boghossian [1989] for general discussion). Now, Millikan [1990] argues, as one would expect, that the meanings of utterances and thoughts are indeed not given by any disposition but rather by the biological purposes that are grounded in the histories of the relevant cognitive systems. Roughly, it is the contribution such a representation-producing system has made, however indirectly, to the reproductive success of hosting organisms, that determines the meaning of the representations they produce.

[358] In later work, Millikan calls such substance templates 'real categories'—see Millikan [n.d., ch. 1, §8].

of concepts with essences. Instead, the underlying metaphysical story should explain why the cognitive economy works. Millikan writes:

> Not just any possible world could be thought about or talked about. A known or thought-about world must be organized in a suitable way. [Millikan n.d., ch. 1, §2][†]

Thus, the underlying metaphysics is settled, at least in part, by an inference to the best explanation: the best explanation for the fact that we have the cognitive and linguistic abilities Millikan has identified is that the world has a certain underlying structure.[359] Reality-for-us is cleanly separated from reality-itself: there is no need for our unicepts to capture the 'true ontological grounds' of the substances they track (except, of course, for the minimal requirement that is embodied in the substance templates).

Millikan introduces the kind of ordering she takes our world to exhibit by way of the following picture [see ibid, ch. 1, §2]. Imagine a huge, multidimensional graph, whose dimensions represent different property continua—one for mass, one for volume, one for temperature, etc.[360] Imagine that this multidimensional graph in effect represents a huge logical space that allows for every relevant aspect of physical objects to be represented by positions along its various dimensions. The result is that every physical object, as it is at a given time, occupies a distinct dot in that space. Now, Millikan observes:

> First, when every object there actually is has been represented by a properly located dot on this graph, all but vanishingly small areas of the graph will remain empty. That is because nearly every arbitrary list of logically compatible properties that would exhaustively describe an object ... fails to describe any real object. Second, where not empty the graph would contain mostly clusters of dots that were in close proximity along multiple dimensions, closely clustered in large subspaces of property space. [Millikan n.d., ch. 1, §2]

The clusters correspond to Millikan's substances: they are what make language and thought possible because they are what unicepts can latch onto. Had the logical space been filled uniformly, such that rabbits shaded off gradually into Gothic cathedrals

[†] *Here, and below, we cite, with kind permission from Ruth Millikan, from a draft of the first chapter of* Unicepts, Language, and Natural Information*, which is to be published at Oxford University Press.*

[359] We discussed the link between such a metaphysical inference to the best explanation and Modern picture-based metaphysical realism in §1.1.1, p. 24 and §1.2, p. 38.

[360] Our use of 'graph' here has no connections to graph theory—we adopt the label from Millikan. Notice that many property continua presumably cannot be thus combined into one multidimensional space—as Elder [2011, §§5.4–5.5, 2013b] argues in a different setting. However, that does not matter for the point of the image.

and CD's shaded off gradually into McDonald's restaurant buildings, there would be no such clusters for unicepts to latch onto.

Now, the mere *factual* clustering of things in this way is not enough for Millikan's purposes. She writes:

> I will be using the name "real kinds" for clusters that are formed within any general category of things, within any multi-dimensional logical subspace, so long as this clustering occurs *for a univocal reason*. [Millikan n.d., ch. 1, §2; emphasis added]

The point is that a clustering that just happens to occur, against all probability, by some kind of cosmic fluke, in no way guarantees its own continued existence in the future, and hence is not something that unicepts should latch onto—there is nothing to be learned 'from one encounter of what to expect on other encounters'. That is, induction works only if there is such a 'reason' for the clustering.

Relatively straightforward cases of clustering 'for a reason' include what Millikan calls 'eternal kinds':

> Certainly the elements and chemical compounds each display a large and characteristic set of dispositions and properties that are collected together and quite neatly separated from the properties of other such stuffs. ... The reason for their property clusterings ... [depends] on the strictly limited variety of elements of which the physical world is composed and upon the patterns that are natural law. [ibid, ch. 1, §6]

That is, eternal kinds are second-gear natural kinds, in our terminology, governed by second-gear natural laws. Millikan here adopts a second-gear metaphysical picture; she would reject first-gear reductions of such basic second-gear notions.[361]

More interesting for our purposes, however, are what Millikan calls 'historical kinds':

> An historical kind bundles a set of properties together repeatedly not merely owing to lawful relations among these properties but owing to a common historical origin or source. These members are alike partly because causally connected, ultimately, to the same individual things. [ibid, ch. 1, §4]

The thought is that the particular clusterings in our logical space that are historical kinds are not necessitated by natural law, as the eternal kinds are. There could have

[361] A quote exhibiting her stance on this matter can be found in Millikan [1999, p. 47]: 'I am going to bypass the traditional Hume-inspired question about the difference between true but accidental universal empirical generalizations and genuine natural laws, and simply assume there is such a distinction, in nature, the reality of which is unaffected by the fallibility of our epistemic capacities to track it.'

been others instead.[362] Still, there is a reason for such clusterings: they rest on cyclical or repeating processes of production and reproduction:

> Continued and/or duplicated production of a uniform type will often depend on natural selection or on some likeness of natural selection, viewing natural selection here . . . as accounting for stability in many characteristics of kind members over time and space. Natural selection or selection for artifact reproduction accounts for clumping, first, by throwing out the junk that would otherwise accumulate through mutation or error and, second, by preserving materials in that process for making more that are like the originals rather than like other things. . . . Wherever a stabilizing process of this sort happens to settle itself into the world, it uses up materials and locations in making its proprietary continuants or duplicates, removing these materials from availability for the constitution of other, hence different, things. Once randomly started, stabilizing processes tend to take over. [Millikan n.d., ch. 1, §3]

Historical kinds comprise not only biological species, but also their parts and stages and artifacts. The entire cultural and social realm of traditions, games, furniture, dances, governments, monetary systems, universities, etc. also fall under the heading of historical kinds, according to Millikan. The category may also includes such entities as rivers, planets, rocks, stars, galaxies and the like.

Now, interestingly, Millikan takes individual objects to be very similar to historical kinds, in key respects:

> Individual objects we have names for have persisted. Their structures have managed to maintain or reconstitute themselves . . . . Not just any physical configuration can do this. The configuration of matter that makes up a stage of an explosion does not endure through time. But if a structure has the right form to endure or to keep reconstituting itself over a period of time, as does the structure of a rock or of a living organism, so long as it stays around it is precluding its materials and its location from belonging to any thing of an opposing kind. Having once arrived, enduring individuals, like historical kinds, tend among them to use up much of the world. [ibid, ch. 1, §5]

What distinguishes individuals from historical kinds is that the latter but not the former can be spatially discontinuous—though, as Millikan says, 'how exactly to distinguish continuing individuals from historical kinds seems ultimately a verbal issue' [ibid, ch. 1, §5].[363]

[362] 'Given only natural law', Millikan writes, 'other species might have inhabited the world in their place; other stably recurrent combinations of properties might have been there instead.' [Millikan n.d., ch. 1, §4].

[363] Elder [2013a] critically discusses Millikan's underlying understanding of persistence: it allows for something to belong to many historical kinds at once (e.g., human being, human adolescent, diabetes sufferer, etc.), and hence provides no clear answer as to whether or not the thing in question persists through certain changes (the human being, but not the adolescent, persists through growing older—how can that be if they are one and the same?). We will not discuss this aspect of Millikan's view here. However,

It becomes apparent, as we get to know the kind of metaphysical picture that Millikan develops, that the realm of historical individuals and kinds is, in an important sense, largely due to mere happenstance: it just so happens that a certain kind of 'stabilizing process' kicks in at a certain point, leading to uniformity in certain respects in what it keeps on producing. But *given* that just that process started, it is *no accident* that the corresponding individual(s) are alike in just those respects. It is important to stress that this non-accidentality holds for the similarity amongst members of the same animal species as well as for the similarity between successive stages of a rock (we will use this fact in our critical examination of Millikan's reduction below). Relatedly, it is important to realize that there is no involvement in the 'purposive and nonaccidental order', up until now: the stabilizing processes Millikan has in mind are purely causal, second-gear processes that are stable either because of natural law alone or because of natural law in combination with a stable background setting (a relatively stable environment on an enduring planet, say).[364]

So, how *do* we get from this understanding of historical kinds and individuals to the 'purposive and nonaccidental order'? Here, we have to go back to the detailed exposition of the notion of 'proper function' Millikan developed in her first book. There, she writes:

> [I]t is the "proper function" of a thing that puts it in a biological category, and this has to do not with its powers but with its history. Having a proper function is a matter of having been "designed to" or of being "supposed to" (impersonal) perform a certain function. The task of the theory of proper functions is to define this sense of "designed to" or "supposed to" in naturalist, nonnormative, and nonmysterious terms. [Millikan 1984, p. 17]

Millikan is thus very clear about her ambitions: living things are things that have proper functions, and a successful theory of proper functions is a second-gear one. Now, Millikan's strategy is as follows:

> A device that has a *direct* proper function has this function as a member of a special kind of family that I will call a "reproductively established family." Things similar to one another form a reproductively established family, in the simplest cases, when these things are similar to one another because something like copying has been going on. [ibid, p. 18]

Reproductively established families are families of things that are alike in certain respects because later ones are copied from earlier ones in those respects. The copying

we should point out, first, the similarity between this issue and the issue of species development to be discussed below (§7.4.2, p. 307), and second, the similarity between Millikan's view on individuals as close to historical kinds with Ghiselin's view of biological species as close to individuals (see §7.4.2, p. 303).

[364] Perhaps stability purely because of natural law, without the help of any contingent background facts, does not exist. However, that doesn't matter for our purposes.

in question, Millikan urges, should be explainable entirely in terms of 'natural law':

> For *B* to be a reproduction of *A* it is necessary only that there be *some* way of describing *B*'s causal history, holding certain conditions constant (mentioning these as initial conditions or simply as conditions that as a matter of fact entered the scene at this or that point in *B*'s history) such that it is explained why *B* had to be *like A*, whatever the character of *A*, within a certain range of character variation. [Millikan 1984, p. 20]

We could say that, given the relevant circumstances, a causal mechanism explains why *B* is like *A* in certain respects—we are still within the second conceptual gear. And, indeed, reproduction in this sense is only the first step towards a theory of proper function, it is not sufficient for the items involved to have any proper functions, of course.

Millikan illustrates her technical definitions of the relevant concepts (reproduction, reproductively established family) with the help of examples. One example is a copying machine: a purely second-gear description of the copying machine and of the black-and-white pattern on the original *A*, together with the fact that a certain force is exerted ('as a matter of fact') on the 'copy'-button, results in *B*, the copy, being like *A* with respect to just that black-and-white pattern. Another example, illustrating the fact that the same thing can be a copy of multiple things at once, concerns a silver cross that also resembles a vine of olive leaves: it is copied, with respect to overall shape, from earlier examples of Christian crosses, while also being copied from the olive leaves (or pictures thereof) that were studied by the artist who produced it. Now, Millikan remarks:

> The laws operating in situ here are derived in part from the structure and state of the individual nervous system of the artist, just as, in the case of the copying machine, the laws are derived in part from the structure of the insides of the machine. [ibid, p. 21]

Here already, at this early stage of our investigation into Millikan's reduction of third-gear to second-gear thought, we notice that a problem arises. For it is a mere article of faith to think that the 'structure and state of the individual nervous system of the artist' should, just as the structure and state of the copying machine, enable a second-gear explanation of the relevant properties of the silver cross in the way Millikan thinks. If our considerations in §7.2 are right, the second-gear indeterminism that is a precondition for the existence of third-gear organisms may very well leave open, to a considerable degree, what will happen at the relevant moment (shortly before the production of the cross). But we should not get ahead of things—we will come back to the issue below (§7.3.3).

Now, given this understanding of copying or reproduction, Millikan defines 'first-order reproductively established family' as follows:

> Any set of entities having the same or similar reproductively established characters derived by repetitive reproductions from the same character of the same model or models form a *first-order reproductively established family*. [Millikan 1984, p. 23]

Millikan warns us that this definition only provides the basis for what she is really after. The reason is that hearts, for instance, are not copies *one of the other*. But, still, they are the results, ultimately, of things that *do* belong to first-order reproductively established families, such as genes (or so the thought goes). In particular, their production is explained by members of such a first-order reproductively established family performing their direct proper function(s).[365]

Let us move on, now, to such direct proper functions—the crucial bit of Millikan's theory. She writes:

> Where m is a member of a reproductively established family R and R has the reproductively established . . . character C, m has the function F as a direct proper function iff:
>
> (1) Certain ancestors of m performed F.
> (2) In part because there existed a direct causal connection between having the character C and performance of the function F in the case of these ancestors of m, C correlated positively with F over a certain set of items S which included these ancestors and other things not having C.
> (3) One among the legitimate explanations that can be given of the fact that m exists makes reference to the fact that C correlated positively with F over S, either directly causing reproduction of m or explaining why R was proliferated and hence why m exists. [ibid, p. 28]

Members of a reproductively established family R are alike in certain respects, that is, they share a certain character C. That character presumably has many causal consequences—to cite a classic example, hearts not only pump blood but also make sounds, push against neighboring tissues, etc. Only some of these causal consequences will play a role in an *explanation* of why later members of R came into existence: these are the direct proper functions of the members of R. And the conditions under which the coming about of these special causal consequences was relevant to that explanation are the normal conditions for the performance of that

[365] The structure of Millikan's definitions is, then, a recursive one: first-order reproductively established families have proper functions that may give rise to higher-order reproductively established families that then also have direct proper functions. The direct proper functions of the latter depend on those of the former, but not *vice versa*—there is no threat of circularity (thus Plantinga [1993, p. 203] is wrong when he thinks there is such a threat).

function.[366] If all this holds, so Millikan, we may say that current members of *R* *should* have those special causal consequences when the relevant normal conditions obtain: the normativity of the third conceptual gear has just entered the scene.

Millikan's story is decorated with many more bells and whistles, which play important roles for her views on cognition and language: as we noted, she adds higher-order reproductively established families, and then also derived proper functions, stabilizing and standardizing proper functions, etc.[367] For us, however, the essential core of the theory suffices: the direct proper functions of (parts or aspects of) members of first-order reproductively established families. That is the basis for the more complicated recursive definitions. It is here that, according to Millikan, biological normativity and teleology arise out of purely second-gear goings-on. Let us move on, therefore, to a more critical assessment of this reductive understanding of third-gear thought.

7.3.3 Against Historical Reductions

It is important to stress the consequences of that most salient aspect of Millikan's view: that a thing's functions supervene on its history, and do not depend on its actual make-up. That means that we can separate the purely second-gear, causal order from the 'purposive and non-accidental order' quite strictly: if we consider an organism (say) only as it is *now*, no third-gear concepts can be applied. Present configuration is a second-gear matter; history provides the grounds for using third-gear concepts (if any). We noted this above, where we presented the illustrations of the copying machine and the artist (on p. 288). One notorious case illustrating this aspect of Millikan's theory is Davidson's 'Swampman':

> Suppose lightning strikes a dead tree in a swamp; I am standing nearby. My body is reduced to its elements, while entirely by coincidence (and out of different molecules) the tree is turned into my physical replica. My replica, The Swampman, moves exactly as I did; according to its nature it departs the swamp, encounters and seems to recognize my friends, and appears to return their greet-

[366] Millikan calls those conditions Normal conditions, and the resulting explanation a Normal explanation—capitalized to stress the fact that it is not some statistical notion of normality that is involved (i.e., 'in most cases', or 'on average', or the like) but rather a teleological or 'quasi-normative' one. See Millikan [1984, p. 5].

[367] Something has a derived proper function *F* if its production by some other thing is a means for that other thing to perform *F* [see Millikan 1984, chs. 1–2, and also Millikan 1989]. Unicepts, then, have derived proper functions: they are produced, according to Millikan, by a mechanism that has as its proper function to produce such items in order for the organism to track various substances. That is why even the very first unicept of, say, bacteria that was ever developed (in Antonie van Leeuwenhoek in 1683) had as its proper function to track bacteria—despite having no ancestors with that function.

> ings in English. It moves into my house and seems to write articles on radical interpretation. No one can tell the difference. [Davidson 1987, p. 443]

Now, Davidson of course has his own reasons to deny that Swampman can have thoughts, recognize and greet friends, and the like. For us, however, the case serves as an illustration of the strict division that Millikan imposes on the second- and third-gear aspects of living things. For Swampman does not have the history that could provide it with any kind of functions. Hence, according to Millikan, it is no human being, it does not have a heart or kidneys or a brain—it is in no way a living thing, *even though, as Davidson writes, no one can tell the difference*.[368]

It is perhaps good to contrast Millikan's response to this case with our own. The Swampman case is, on our view, most likely to be an impossible scenario. We do not mean that the production of the replica is somehow impossible (though perhaps it is), but rather that *if* the replica is produced, it cannot be something that is organized by the life-form of human beings and hence will disintegrate fairly quickly—very similar to the way corpses do, presumably. There will be no strange physical automaton walking around that is indistinguishable from Davidson; rather, there will be something that looks like Davidson's corpse lying in the swamp. The reason is twofold. First, the replica cannot be an instantiation of the life-form 'human being', since human beings come of human beings and by nature develop in a certain way (i.e., they grow up), which is not what goes on in the present case. And second, life-forms make a significant contribution to the persistence of an organism (as we argued in §7.2), so if no life-form is present, the matter that is supposed to constitute Swampman will go its own way—as it does when organisms die.[369]

In effect, then, Millikan's view tells us that the central notions of the third conceptual gear, or, in her own terms, of the purposive and nonaccidental order, summarize complicated historical-causal explanations for the proliferation of second-gear things sharing certain characteristic features. Put less friendly, such notions embody, on Millikan's view, a promissory note to the effect that such second-gear explanations

[368] Millikan explicitly discusses Swampman-like cases at several occasions. See, e.g., Millikan [1989, pp. 292–3, 1996, 2010, pp. 76–9].

[369] Might it not be that the life-form of human beings *does* allow for a deviant form of generation, omitting part of the normal developmental cycle? That is hard to believe, but still, even if it does, it is clear that a mere *physical* replica of Davidson's body is then insufficient: minimally, all the life-processes going on would have to be replicated as well. And here lies the problem: life-processes cannot occur independently of the organism to which they belong. That is to say, one cannot generate an organism *by* creating life-processes, at best one can create life-processes *by* creating an organism. When the first living things emerged, the two perhaps came together—generation of a (very simple) organism and generation of the corresponding life-processes. But human beings are not such organisms. The idea of spontaneous generation *of human beings* can be safely rejected. (See also our brief remarks on the origin of life in §7.4.1, p. 299 below.)

can in principle be given (which signifies the second-gear metaphysical dogma). Strangely, then, whether a thing has any proper functions makes no difference at all, strictly speaking, to what it does—the functions concern its history only; they explain, in a certain way, why the thing has the second-gear, non-normative powers it has. Millikan's response to Swampman makes that clear: there is no difference between Davidson and Swampman except in their histories. And these histories are completely inert—what makes Davidson and Swampman do whatever it is they do is their second-gear make-up, not their histories, so if that second-gear make-up is identical, 'no one can tell the difference'. Thus, when Millikan says that members of historical kinds are alike 'for a reason', this reason can only be a second-gear reason—there is no clustering for third-gear reasons. Her reduction of third-gear notions seems to be complete.

Earlier, we found that reductions from higher to lower gears tend to result in reversing the order of explanation involved (e.g., in our discussion of laws in §5.2.2, on p. 171, and in our discussion of causation in §6.1, on pp. 184 and 190). Here, we may observe another instance of this phenomenon. Instead of explaining the fact that biological items do certain things by reference to their purposes or functions, Millikan's view implies that we should explain their having those purposes or functions by reference to the fact that they (or, rather, their ancestors) do the relevant things. However, since reduction to second-gear thought is much more broadly accepted than reduction to first-gear thought, in part because acceptance of the latter entails a form of acceptance of the former, this observation will not be all that convincing, of course.

Now, if our claim as to the irreducibility of third-gear thought is correct, we should be able to point out where exactly Millikan's account comes into trouble. There are two options:

(1) The account fails to yield a truly third-gear concept of proper function, hence fails to distinguish the living from the non-living, contrary to Millikan's claim.

(2) The sort of explanation that is involved in grounding direct proper functions relies on third-gear notions already, again contrary to Millikan's claim.

These are, as was to be expected, versions of what we encountered before: Thompson's sub-metaphysical Scylla and tautological Charybdis (see 7.1). It is, of course, possible to read Millikan's view in such a way as to fall under (2): that would be the case if, say, the normal explanation for the proliferation of the reproductively established family in question had to be a third-gear explanation that concerns in-

stantiations of a life-form rather than second-gear physical objects. Given Millikan's reductive intentions, however, it is clear that we should prefer not to read her in that way. It turns out, however, that then (1) applies instead, as we will now show.

Consider a particular rock r. According to Millikan, rocks exist because a 'stabilizing process' started, at some point in the history of Earth. That is, given the environment on the surface of our planet, rocks tend to persist stably. That is not necessitated by natural law alone, of course: our rock r would, when placed in the interior of the planet, dissolve. Hence r is a 'historical individual': its existence can be traced back to a certain contingent origin involving relevantly stable background conditions. Now, Millikan takes the internal structure of r a moment ago to be a cause of its present internal structure (and we will assume for the moment that this is correct). If its internal structure had been different, a moment ago, it would have a different internal structure at present. Thus, we may classify the successive stages of our rock r_1, r_2, etc. as members of a first-order reproductively established family: they are the result of a kind of copying mechanism (see the definition of reproduction quoted on p. 288 above). Let us say that r_N is the present stage of r. Now, let us assume that certain aspects of r's make-up, which we call C, account for its not being very sensitive to erosive forces such as wind and water. This effect of C we call F. The stages $r_1, r_2, \ldots, r_N$ all have C, hence they all perform F. C, then, is part of the reproductively established character of the stages $r_1, r_2, \ldots, r_N$. Let us further assume that there are variants of r, in terms of chemical constitution and internal structure, that lack C and are, as a result, much more sensitive to erosion. Together with r, they form a set of items S. Then, we can see that the definition of direct proper function applies (see p. 289 above):

(1) Certain ancestors of r_N performed F.

(2) In part because there existed a direct causal connection between having the character C and performance of the function F in the case of these ancestors of r_N, C correlated positively with F over a certain set of items S which included these ancestors and other things not having C.

(3) One among the legitimate explanations that can be given of the fact that r_N exists makes reference to the fact that C correlated positively with F over S, either directly causing reproduction of r_N or explaining why r was proliferated and hence why r_N exists. [Millikan 1984, p. 28]

It seems, then, that stages of rocks have, as a direct proper function, to resist erosion. In general, the persistence of certain types of rocks over larger periods of time, while

other types wither away, resulting in such beautiful rock formations as can be found on many places on our planet, may be viewed as the result of a kind of natural selection process. After all, if there were to be rocks that were *even better* at resisting erosion, because they possessed some suitable variant of *C*, they would outlast our original rock *r*. Compare Millikan's following remark:

> Putting things intuitively, products of evolution . . . are reproduced or continue to be proliferated because they, *rather than certain other things*, have been associated with certain functions. If certain other things had correlated better with these functions, the chances are these other things would have been reproduced or would have proliferated instead. [Millikan 1984, p. 27]

Reflecting on this example, we arrive at the suspicion that it generalizes quickly: virtually the entire realm of historical kinds and individuals will then be endowed with direct proper functions in Millikan's sense. As long as there is some room for things to have been different (which is the case for historical but not for eternal kinds), and as long as there is some *de facto* explanation as to why a certain stabilizing process rather than other stabilizing processes was initiated, there is bound to be *some* sense in which the relevant thing or kind is the result of a kind of natural selection process. Millikan's 'purposive and nonaccidental order' turns out to be much more encompassing than initially thought.

We should be careful, now, in drawing lessons from this perhaps unexpected application of Millikan's notion of direct proper function. Millikan explicitly says that her definitions are meant as 'theoretical definitions', not as capturing *our use* of terms like 'function' and 'purpose' [see Millikan 1989, p. 291]. Indeed, she usually is no fan of conceptual analysis—if empirical findings or theoretical considerations suggest that reality is better understood by using certain notions differently from the way we ordinarily use them, then so much the worse for our ordinary use. Hence it is open to her to simply accept that rock stages have proper functions (and if not these, for some reason, then surely there will be other such cases). We cannot claim to have 'refuted' Millikan by way of *reductio ad absurdum*.[370]

[370] That is what Alvin Plantinga seems to be doing in his criticisms of Millikan—see Plantinga [1993, pp. 203–4]. He writes that '[t]he logical relations between some of the notions aren't at all clear to me' [p. 203], and indeed gravely misunderstands the structure of Millikan's carefully developed chain of definitions, which becomes especially apparent in the putative problem cases he discusses—e.g., cases involving artifacts without ancestors, which of course have derived rather than direct proper functions according to Millikan (derived, that is, from the maker's intentions). On the other hand, Michael Rea [2002, pp. 118–22] comes somewhat closer to our sort of criticism. He uses the fictitious example of reproductive clay crystals made famous by Dawkins [1996, pp. 150–4]—but then adds further cases involving demigods and genetical engineering which suffer from the same defect as Plantinga's cases. Rea concludes that '[i]t simply does not follow from the fact that a thing has a certain kind of natural causal history that it is in

Still, as we observed, Millikan's project was to argue that 'it is the "proper function" of a thing that puts it in a biological category' [Millikan 1984, p. 17]—hence we may question what is left of that claim once we see that proper functions do not demarcate biological categories but rather her own much broader category of historical kinds and individuals. It seems that a modification of the claim is necessary. Consider, then, a modified, Millikan-based view on life which says: anything having proper functions in Millikan's sense *and also having further characteristics C* belongs to a biological category. Then, obviously, we are back where we started—with Thompson's criticism of a list-of-life. The lesson we draw from our observations, then, is that Millikan's proposal in the end indeed turns out to fall under Thompson's 'sub-metaphysical Scylla'-category: though much, much more sophisticated than the mere remark that living things 'are highly organized' or 'contain DNA' (the examples we discussed in §7.1 above), Millikan's account still boils down to a proposal of the same kind: a purely second-gear feature is singled out as being definitive—well, not directly of life, but nonetheless of her 'purposive and nonaccidental order', which turned out to comprise all historical kinds and individuals. Of course, the second-gear processes on which Millikan's view rest, such as processes of reproduction, *are* special—unlike, say, processes of evaporation and incineration. But, as we remarked, following Thompson, they are special in that they play a special role *for living beings*—a remark best put under the heading of Thompson's 'tautological Charybdis'. *As such*, however, the features that play a role in Millikan's definitions are just not at home in the third conceptual gear, as we have illustrated. Millikanian functions are second-gear simulations of true, third-gear functions—just as perdurance is a first-gear simulation of true, second-gear persistence (see §6.2.1).[371]

Now that we have critically reflected on Millikan's reduction of biological normativity, there is another observation we should make, precisely concerning the putative normativity involved in her notions of direct proper function and the like.

---

some metaphysically important sense supposed to do whatever it is that its causal history has determined it to do.' And that is, in a way, what we are concluding too—but only if we read 'metaphysically important sense' to point towards third-gear thought.

[371] Might there not be amendments to Millikan's position that make it immune to arguments of the kind we presented? For instance, one might hold that mere copying, which is what Millikan's theory of function rests upon, will not do: it should be replaced by a notion of reproduction that is genuinely biological in nature. Something along the lines, perhaps, of Griesemer [2000]: 'reproduction is multiplication with material overlap of products with the capacity to develop. Since development is the acquisition of the capacity to reproduce, we can say that reproduction is progeneration [i.e., multiplication with material overlap] of entities with the capacity to develop the capacity to reproduce.' Some reflection, however, makes clear that this *either* merely complicates the definition and hence the counterexamples, *or* already incorporates third-gear categories (i.e., in the assumed understanding of 'capacity', 'development' etc.)—that is, we are led back to Thompson's submetaphysical Scylla and tautological Charybdis once more.

Once we see what exactly that normativity is grounded in—historical explanations of proliferation—we see that it in fact consists in no more than a certain type of justification: given that the coming about of A is explained, in part, by reference to its ancestors' performing a certain function F, under certain conditions, and in a certain minimal proportion of cases relevant to the explanation, *we may infer* that A will also perform F under those conditions, with the relevant probability. If it turns out that A does not do so, A fails to perform its proper function—so Millikan says, using normative terminology. But in fact the norm is a reification of the justification involved: contrary to what we held ourselves to justifiably believe, A does not do F. Apart from the epistemic realm, there is no norm—just a certain causal history. Millikan may, of course, accept this conclusion: she wanted to have a 'nonnormative' theory of biological normativity, after all. For us, however, it is clear that this in effect drives a (Modern picture) wedge between the third conceptual gear and reality as it is in itself—third-gear teleology and normativity turn out to be a feature of our cognition (or of reality-for-us), not of reality (reality-itself).

This concludes our discussion of Millikan's views. One final remark is in order, however. It is worth reminding ourselves that the entire discussion was meant to illustrate our general line of thought: third-gear thought is irreducible to second-gear thought, hence third-gear phenomena stand on their own. If that is right, denying this line of thought leads to a failure to grasp and understand the realm of the living—Thompson's 'sub-metaphysical Scylla'. We have seen that this is what happens in the case of Millikan's reductive view. Of course, if we simply forget about the reductive understanding of third-gear categories that Millikan proposes, and take her use of third-gear thought at face value, we may appreciate the way she employs it for various philosophical purposes—but then we arrive at Thompson's 'tautological Charybdis'. Now, *what we have just said only makes sense if we accept the possibility that third-gear thought comprises a realm of its own*—as we claim it does. Insofar as it *does* make sense, then, it illustrates our reasons for making that claim.

7.4 Evolution and Life-Forms

'Nothing in biology makes sense except in the light of evolution', as Theodosius Dobzhansky, a major figure in the development of 20$^{\text{th}}$ century biology, quipped.[372]

[372] Thus reads the title of Dobzhansky [1973]. In that article, Dobzhansky's aim was to stress the status and importance of evolutionary theory in contemporary biology in the face of strong anti-evolutionist campaigns of religious origin. That particular controversy is irrelevant to our concerns.

The thought is that, without the idea of evolution, biology is just the inventory of the kinds of organisms that happen to be around—a huge classificatory scheme of a Linnaean kind—plus factual descriptions of the workings of ecosystems, individual organisms, and their organs and parts. With the thought that life evolves, however, this entire static scheme starts to make sense against a dynamic background: we understand, in general terms, how the kinds of organisms we find around us could have come into being, how the impressive specialization we observe could have come about, why some species are more alike than others—in short, instead of merely noting *that* the organisms around us are the way they are we come to understand *why* that is the case. It all makes sense in the light of evolution.

Now, enlightening though the thought that life evolves may be on this score, it has also been taken to imply a very radical departure from the pre-Darwinian understanding of life on different scores. In particular, it has been taken to support a reductive understanding of life, and it has been taken to imply the untenability of traditional, Aristotelian essentialism about species.

We could be very brief on both issues. First, evolution does not all by itself imply reduction—there is nothing wrong with the thought that third-gear organisms are the subjects of evolutionary processes—hence our anti-reductionist, third-gear understanding of life is not threatened by it. And second, essentialism is quite compatible with evolution: that earlier life-forms give rise to later life-forms through a gradual process of evolution is perfectly intelligible, especially if one takes into account the flexibility of living beings—their capacity to adapt themselves in remarkably diverse ways to different circumstances. However, such brevity is unlikely to be very illuminating—it will be more helpful for our purposes to look into both issues a bit more carefully.

With regard to the first issue, the relation between evolutionary theory and reductionism, we may observe that our discussions so far, and in particular our reflections on Millikan's reductive views in the previous section, have led us to the conclusion that reductionism about life, about the third conceptual gear, is not to be recommended. Yet the neo-Darwinian picture of evolution as being due to random variations feeding a 'mechanism' of natural selection is very close to the picture of a 'blind watchmaker', to use Dawkins's metaphor [Dawkins 1996]—it invites the reductive thought we aim to dismiss: that teleology can be 'analyzed away' by using evolutionary theory as a reductive theory. Given that we reject this application of evolutionary theory, it may be helpful to clarify our positive understanding of it. That is our task in §7.4.1 below.

We then turn to the second issue, which is related to the so-called 'species problem', in §7.4.2. There, we argue that the issue in fact rests upon a terrible confusion over what a species is, and in particular on a confusion of epistemological issues with metaphysical ones. Keeping in mind the peculiarities of third-gear life-forms, as opposed to second-gear natural kinds, turns out to be quite helpful for dissolving the confusion.

7.4.1 Evolution and Reduction

The idea of natural selection can be taken as a purely second-gear idea: a given range of (second-gear) background circumstances will favor the persistence (or recurrence) of some kinds of objects and stuffs over others—as in our application of Millikan's natural selection-based account of proper function to rock stages (see p. 293 above). As such, we concluded that it does not give rise to any interesting (third-gear) notion of function—at best, we arrive at a second-gear simulation of that notion that captures contingent stability of a sort that fails to capture what is distinctive of life (see §7.3.3).

Of course, our results with regard to such a reductive account do not show that natural selection, on this second-gear understanding, has no role to play in our understanding of animate nature. If we drop the reductionism, we can still say that natural selection determines in what way life-forms appear on the scene. We are then talking about natural selection *as it operates on living organisms*. Thus understood, the idea of natural selection presupposes rather than grounds third-gear concepts.

It seems, then, that evolutionary theory and reductionism do not form a package deal. And we are not making a very contentious claim in drawing this conclusion. To illustrate, consider what Stephen Gould writes in his voluminous book on the 'Structure of Evolutionary Theory':

> As functionalist theories, both Lamarckian soft inheritance and Darwinian natural selection share a defining premise that environmental information about adaptive design somehow passes to organisms, and that organisms then respond by fashioning traits to enhance their competitive ability within these environments. [Gould 2002, p. 1179]

That is to say, natural selection, as understood by Darwin, indeed should be understood to be natural selection *as it applies to organisms*, not natural selection *as it gives rise to organisms*.[373] Evolutionary theory, thus understood, is neutral with regard to the issue of reduction.

[373] 'As it applies to organisms' should not be understood to include a specific stance on the debate over what the 'unit of selection' is—individual organisms, species, genes, or something else. For that debate, see, e.g., Sapienza [2010] and Burian [2010].

This neutral stance seems to be widely endorsed also within the philosophy of biology. To illustrate, we may point at the first pair of chapters in the recent *Contemporary Debates* volume on the philosophy of biology [Ayala and Arp 2010, Part I], which is devoted to the question whether it is 'possible to reduce biological explanations to explanations in chemistry and physics'. Evelyn Fox Keller, writer of the first chapter, defends an affirmative answer to that question; yet she writes:

> Natural selection—at least since the neo-Darwinian synthesis and, probably, ever since Darwin—has been conventionally understood as requiring the prior existence of stable, autonomous, and self-reproducing entities.[374] [Keller 2010, p. 23]

Again, it is clear that we should, thus, distinguish evolutionary theory from the reductive project—evolutionary theory *as such* need not be a reductive theory. It simply tells us something illuminating about 'stable, autonomous, and self-reproducing entities', that is, about third-gear organisms—namely, that instances of new life-forms come to be through a preceding process of evolution acting on pre-existing organisms.

The need to distinguish evolutionary theory from the reductive project becomes particularly salient when considering the question how life came into being. On a reductive view, we are to imagine an exclusively second-gear world within which life somehow emerged without adding anything not explicable in second-gear terms. That is, the origination of life has to be explained without remainder in second-gear terms; a Millikan-like story will have to be told concerning the emergence of functions and purposes out of such second-gear ingredients. Our critical discussion of such a theory above indeed underlines (and extends) the observation Gould and Keller make: when the notion of natural selection is divorced from its biological context in order to provide a reductive basis for a theory of function or purpose, the resulting notions no longer separate the living from the non-living. That is exactly what we found in the case of Millikan's views. Natural selection, as a phenomenon within the realm of the living, presupposes third-gear concepts.

On a non-reductive view, by contrast, we are to imagine a world that includes the possibility of third-gear phenomena but does not yet manifest any of them. The

[374] Interestingly, Keller and Dupré, the writers of these two chapters, which are supposed to defend opposing answers to the given question, both notice that their views do not seem to be all that distinct. A clear sign that there is something odd about the reductionism debate. It is noteworthy that both Keller and Dupré find it important to stress that they are 'materialists' [Keller 2010, p. 19, Dupré 2010, p. 33]. Keller has it relatively easy, for she can unproblematically state that '[a]s a materialist, I am committed to the position that all biological phenomena, including evolution, require nothing more than the workings of physics and chemistry' [p. 21]. Materialism, for Keller, is second-gear metaphysics. Dupré, however, has a harder time stating his position. In our interpretation, he tries to reject second-gear metaphysics, yet fails to identify what exactly that amounts to—his endorsement of 'materialism' stands in the way of a successful move towards third-gear anti-reductionism.

emergence of life may then be explained by pointing at suitable (second-gear) circumstances in which primitive life-forms may have come to be instantiated. Thereby, third-gear organisms started to spread, life in all its diversity started to develop—evolution began. The entire story, however, from the first instantiation of a life-form to the colorful plurality of living beings we find around us, is to be understood by using third-gear thought. In such a scenario, evolutionary theory still plays the enlightening role that we mentioned at the start of this section, despite not being employed in a reductive manner.

We may expect, therefore, that most of the (highly speculative) origin-of-life stories that have been proposed as explanations of the emergence of life can in principle be adopted by reductionists and non-reductionists alike—the controversy over reductionism is not about *whether* life could emerge from non-life, but rather, simply, about *what life is*, and hence what such emergence amounts to.[375]

In the end, the question of the origination of life is about what it takes for life-forms to *become instantiated*, and, generally speaking, we are defending an Aristotelian, hylomorphic take on that issue that straightforwardly leads to the kind of answer we have sketched.[376] Anna Marmodoro provides a nicely concise statement of the general idea:

> I understand Aristotle's hylomorphism as a doctrine of *instantiation by change*: the material elements facilitate the instantiation of the essential type by being transformed in accordance with the essential type's organizational principle. When the essence is instantiated, the physical components are all unified into a single thing that is structured and qualified according to the principle the essence stands for—for example, fluids come to be an embryo or a plant, or food comes to be flesh. [Marmodoro 2009, pp. 36–7]

That is to say, in our envisaged primordial event of life-origination the relevant second-gear stuff changes *such as to become a living organism*, which is then supposed to live the life of its life-form—a third-gear life-form has become instantiated. Much like eggs, sugar, butter and flour may change *such as to become a cake*, which is then supposed to taste good—an artifact-form has become instantiated. Here we find ourselves back with that simple principle, the defense of which is what this essay is

[375] Notice, however, that reductionists are committed to such a story of the origin of life in a way that non-reductionists are not. It is open to non-reductionists to follow Louis Pasteur in thinking that 'spontaneous generation' of life is impossible—that, in short, *omne vivum ex vivo*, all life comes from life. See Pasteur's famous public lecture on spontaneous generation from 7 April 1864 [Pasteur 1922–39, vol. II, pp. 328–46]. Pasteur himself took spontaneous generation, if it were true, to support 'materialism'—but then, he thought of such spontaneous generation in terms of 'matter organizing itself', which sounds very much like reductionism. See Roll-Hansen [1979] for a good sketch of Pasteur's attitude.

[376] Recall, in particular, our brief sketch of hylomorphism in §2.4.

devoted to: essences are concepts. Concepts structure reality by being instantiated, and hence they may *come* to structure reality by *becoming* instantiated—life-forms, being concepts from the third conceptual gear, make no exception to this rule; they only bring their specific requirements: suitable conditions for them to be instantiated in. Given that there is at present no such thing as spontaneous generation of life, we may assume that the conditions for instantiation of the life-forms we find around us today depend on pre-existing living beings—sugar maples come from sugar maples, horses come from horses, human beings come from human beings. The primitive life-forms that came to be instantiated when life first emerged have to have been different in this respect.

In short, then, evolutionary theory *as such* poses no threat to the non-reductionist view on life that we are defending. However, we still have to confront the widely accepted thesis that evolution is incompatible, in particular, with the essentialism we defend. It turns out that, *properly understood*, essentialism is under no such threat from evolutionary theory either.

7.4.2 Evolution and Essentialism

The challenge seems to be the following. The traditional conception of living beings as Aristotelian substances that have essences requires the essences involved (the species) to be eternal and unchanging. Darwin's *On the Origin of Species* [1859], however, showed that life develops dynamically, that there always is variation within species on which natural selection can operate, so that species change considerably over time, specialize in diverging ways, and hence give rise to new species in a gradual manner. Contemplation on this picture, so it is thought, reveals that there is something fishy about the essentialist understanding of species. The rigidity that it embodies does not fit with their gradual emergence, and from the viewpoint of evolution, all that matters anyway is concrete populations of individual organisms, not species. In short, it is believed, evolutionary theory has dealt a fatal blow to the Aristotelian, essentialist picture.[377]

Notice, however, that the book which thus supposedly changed biology so radically was, oddly enough, entitled *On the Origin of Species*. Darwin himself, then, was concerned with *explaining* the diversity of species, not with *eliminating* them. It

[377] See, e.g., Hull's early pair of essays entitled 'The effects of essentialism on taxonomy: two thousand years of stasis' [1965a,b]. And, for a more recent source, see Richards, who writes in the introduction to his book on 'the species problem' that 'obviously, evolution implies that species can no longer be regarded as eternal and unchanging' [Richards 2010, p. 2].

does not seem that Darwin was working towards such a radical abandonment of the traditional species concept. So, where does this wide-spread skepticism about species come from?

As it turns out, the story of pre-Darwinian essentialism that was rejected in favor of post-Darwinian anti-essentialism, and thence gave rise to 'the species problem', is just that: a story. John Wilkins, who thoroughly investigated the history of the issue, writes:

> There is a widely held story that most people, if they have thought about the matter at all, usually hold, which has been repeated for 50 years, since the Centenary of the publication of the *Origin*. According to this view, sometimes called the "essentialism story" . . . [378], before Darwin species were held to be universals, classes or natural kinds, which had essential definitions. With Darwin comes recognition of the variation within species, that drives evolution by natural selection. From then on, biologists understood that species were polytypic, that they had no essential properties. Species come to be understood to be biological populations, protected against introgression by reproductive barriers . . . . [Wilkins 2010, p. 141]

Wilkins provides a wealth of historical materials to show that the variation supposedly uncovered by Darwin was, of course, well-known to biologists for centuries, and that the question what exactly biological species are, the 'species question', was also being discussed already for centuries before Darwin. The essentialism that plays a role in the 'essentialism story' is a straw man. It rests on an understanding of essence that was never seriously accepted by any biologist: that a species could be defined by providing a list of necessary and jointly sufficient empirical conditions for being an instance of it.[379] The parallel with the kind of empirical definition of life we rejected in §7.1 should be obvious: the kinds of conditions that would make up such a definition of a species either fail to mark off the relevant life-form (Thompson's 'sub-metaphysical Scylla'; e.g., contains certain genes) or they already contain the very idea of the life-form in question (Thompson's 'tautological Charybdis'; e.g., consists of matter arranged *ant*-wise).

Thus, the 'essentialism' that is supposedly incompatible with evolutionary theory is not the essentialism we have in mind—we hope to make clear, precisely, that having a third-gear essence (i.e., instantiating a life-form) is very different from having a

[378] The label seems to come from Levit and Meister [2006]. Wilkins here also quotes, amongst others, Charles [2002], McOuat [2001], and Winsor [2006] in support of the thesis that what we have, here, is a mere story.

[379] There have been, of course, many *philosophers* who proposed an understanding of natural kinds along these lines, mostly inspired by logical considerations—betraying a first-gear bent of mind. One of the most famous such philosophers is Mill [1884].

second-gear essence (i.e., instantiating a natural kind)—not to mention having a first-gear essence. The sterile construal of essence that was supposedly overthrown after Darwin is, on the face of it, not very sensitive to these differences.

Partly, the source of the confusion over what 'essentialism' amounts to lies in the historical fact that Mendelian genetics became increasingly important in the early twentieth century, after its rediscovery in 1900, leading to the development of population genetics (and ultimately to the 'modern synthesis' of genetics and evolutionary theory in the 1930s and 1940s).[380] This trend led to a quite narrow conception of essentialism:

> Essences according to the new geneticists were ... genetic, and if no groups of organisms shared identical genetic constitutions, then there were no species. [Wilkins 2010, p. 144].

Since it was fairly quickly decided that species do not have such 'genetic essences', essentialism was taken to have been refuted, and the very idea of biological species became suspect.[381] Indeed, the friction between this kind of genetic essentialism and evolutionary theory is easy to see—given that evolution was thought to proceed through natural selection acting on minor changes in the genes, the resulting gradual development from one species out of another is close to incompatible with genetic essentialism. For if *all* genes together determine species identity, every minor change results in a new species (which is not a defensible position), whereas if only *some* genes determine species identity, drawing the relevant line appears to be an unsettlingly conventional affair.[382]

From a metaphysical point of view, it is interesting to note that such skepticism about species has given rise to a rather curious position: the 'individualism' of Ghiselin [1974, 1997]. Ghiselin argues, in effect, that since there are no species in the

[380] The modern synthesis, marked by a milestone publication by Julian Huxley entitled *Evolution: The Modern Synthesis* [1942], involved many of the great evolutionary biologists and geneticists of the time, including Theodosius Dobzhansky, Ronald Fisher, Sewall Wright, J.B.S. Haldane, Bernhard Rensch, and Ernst Mayr. See Gould [2002, Part I, especially ch. 7].

[381] See Wilkins [2009, chs. 5–7] for the relevant background—and see Devitt [2008] for a recent defense of something like genetic essentialism.

[382] One way of escaping this conclusion is by endorsing 'saltational evolution'. Darwin, who supported the thesis *natura non facit saltus* [Darwin 1859, p. 194], was a 'gradualist', thinking that evolution proceeds by accumulation of many small steps [see Gould 2002, pp. 146–55]. But there were (and are) those who believed that some kind of jump-like macro-evolution was necessary to account for the larger differences observed in nature, and that the gradual micro-evolution Darwin championed was thought to give rise to intraspecific variation only. See, e.g., Bateson [1894], and in particular Goldschmidt [1940], who famously named the envisaged results of such larger jumps of nature 'hopeful monsters' [pp. 390–3] (see Gould [2002, ch. 5]). Though ridiculed by the neo-Darwinians, who were establishing the 'modern synthesis' in those years, Goldschmidts ideas have been partly vindicated by later scientists; cases of saltational evolution and even his 'hopeful monsters' have been identified and studied. See, e.g., Theissen [2005], Chouard [2010], and Page *et al.* [2010].

'traditional' sense, they should be understood as scattered individuals that comprise individual organisms as parts, not instances. By conceiving of species as individuals, so the thought seems to go, definitions of the deprecated kind are no longer presupposed—no essential features are required. If at all, then, species may survive the anti-essentialist Darwinian turn in the form of such individuals.[383] Of course, we find this piece of metaphysical speculation to be fundamentally wrong-headed, on by now familiar grounds (see chapter 2): what is the ontologically fundamental category to which such large and scattered 'individuals' are supposed to belong? In other words: what is *their* essence?

Despite all these troubles, both pre- and post-Darwinian biologists 'in the field' appear to have been struggling (and are still struggling) with the very same, serious question: what it takes for some individual organism to belong to one or another species. Put in Aristotelian terms, they have been (and are) interested in 'what-it-is-to-be' a member of that species—which is (roughly) Aristotle's original term of art that came to be translated as 'essence' in Latin. The overall take on the issue appears not to have radically changed since Darwin; it is just that evolutionary and hereditary perspectives have been added to the mix—and have dominated it, of course.[384] According to Wilkins, the core of that mix is roughly as follows:

> If species are indeed thought to have essences, they are of the developmental kind—a lifecycle that reliably generates the morphology, ecological niche occupancy and behavior that is typical. [Wilkins 2010, p. 144]

This is, more or less, what Wilkins calls the 'generative species concept'. Wilkins takes species, as they have been studied throughout the history of biology from Aristotle onwards, to have been understood not as 'logical species' characterized by clear-cut definitions, such as the genetic one, but rather as being 'natural species', which have been thought of along the lines of his generative species concept. Claims about such natural species do not hold strictly invariably but rather 'for the most part' (as

[383] Notice how this conception of species quickly leads to a first-gear conception of persistence (as discussed in §6.2.1 and §6.2.5). That is illustrated by Reydon [2008], who explores the connection between Ghiselin's individualism with the endurantism/perdurantism debate—he writes: 'Species three-dimensionalism and species four-dimensionalism say different things about the nature of species, so asserting that species are individuals without explicating what sort of individuals one has in mind just will not do'.

[384] E.g., Mayr proposed—and refined over the years—what he calls the 'biological species concept'. An early formulation: 'Species are groups of actually or potentially interbreeding natural populations, which are reproductively isolated from other such groups' [Mayr 1940, p. 120]. A later formulation: 'A species is a protected gene pool. It is a Mendelian population that has its own devices (called isolating mechanisms) to protect it from harmful gene flow from other gene pools' [Mayr 1970, p. 13]. As Wilkins notes [2009, p. 189], however, it seems that Mayr is more concerned with the epistemic than with the ontological aspects of being a species.

Aristotle would say).[385] Though not identical to our notion of 'life-forms', Wilkins's generative species concept does come close—a lot closer than the 'essentialism' that was allegedly refuted by Darwin.

It is no overstatement, then, if we say that, if anything, the Darwinian and genetic revolutions have not freed us from an inadequate form of essentialism, but have rather turned our understanding of what life-forms are, which was already a difficult matter, frankly into a conceptual mess. The ongoing discussion amongst biologists and philosophers of biology over the 'species problem', over the question what exactly biological species are, on what their status is, and on what demarcates them—morphological or physiological or genetic characteristics, possibility of interbreeding, evolutionary history, etc.—in fact illustrates this claim. For instance, Mayr already observed in 1942 that there were five different "species concepts" in use amongst biologists [Mayr 1942]; later, Mayden [1997] noted that that list should be expanded to over twenty different species concepts, while Wilkins [2006] extends the list further to 26. Concerning this last list, Wilkins writes:

> Here is a working list of species concepts presently in play. ... for philosophical reasons, I think there is only **one** concept—"species", and all the rest are **conceptions**, or definitions, of that concept. [Wilkins 2006]

Here, we wholeheartedly agree, of course—the one concept is that of life-forms, a basic category of the third conceptual gear; the many conceptions arise from a (perhaps too narrow or exclusive) focus either on certain families of life-forms (e.g., higher animals) or on certain typical distinguishing features of life-forms (e.g., morphological or genetic features). It is an interesting challenge, for the science of biology in general, to tie them all together—it will illustrate how diverse life-forms can be.

From our non-reductive, Aristotelian perspective, we can make three general remarks that may help clear up the conceptual mess. First of all, one should not expect the kinds of definitions familiar from first-gear examples ('logical species') to be available on higher conceptual gears. By contrast, the third-gear essentialism we are defending is an essentialism of *life-forms*, of course, and these, as sketched in §4.4, are not mere lists of empirically detectable, isolated features but rather intricate unities of normative and teleological features, life-processes, and interrelations with other life-forms, all of which inherently refer back to the life-form as a whole.

Secondly, and relatedly, that universality fails in the way it does—that there seem to be no features that are shared by all and only members of a given species—is

[385] See Wilkins [2010, p. 142]. Wilkins also refers to Wittgenstein's notion of 'family resemblance' [p. 145], remarking that family members resemble each other, of course, because of shared generative histories.

typical of third-gear organisms: they can fail to display features they *should* display, and can fail to engage in life-processes they *should* engage in. That is not a problem for species essentialism, it is an interesting feature of it—it illustrates the normativity that pertains to the particularly third-gear form of predication.

And, thirdly, it takes empirical investigation to arrive at the fundamental concepts (essences) of both second-gear objects and third-gear organisms (as we argued in §3.3.3), hence from an epistemological point of view it is not so surprising that the *conceptions* individual biologists have of species in general—and, for that matter, of any given species in particular—is often a mix of *epistemically* relevant indicators and more fundamental aspects.

Reflection on the differences between the third conceptual gear and lower conceptual gears can thus be of help in clearing up the conceptual mess surrounding the species problem. A comparison with a second-gear stuff, like water, may help to illustrate our point. Water appears in different ways: as a gas, a liquid, a solid, in drops, in crystalline forms, in various mixtures such as coffee and mud, etc. Now, the diversity of guises in which we find water does not give rise to skepticism concerning water as a natural kind. Neither, we submit, should the fact of diversity amongst species members give rise to any such skepticism—not even if that diversity is one of gradual and continuous change across the entire population (unlike for water). In the case of water, what matters is that it is *the same stuff* that is present in all the guises. And because second-gear stuff, as we observed, does not *organize* underlying stuff, but simply *is* that stuff, things are relatively straightforward in this case—you can keep track of the same portion of liquid water as it freezes into ice (say). Life-forms, by contrast, do organize the stuff they are made of. Hence *epistemological* questions as to how sameness and difference in life-form is settled are less straightforwardly answered: you can't track the organism by tracking the underlying stuff. It seems that we have to get to know life-forms via their characteristic phenomena, via their specific way of organizing the underlying stuff, their specific way of living. The phenomena in which life-forms are expressed may thus encompass all the different aspects that figure in the biologists' diverse species conceptions: ranging from relevant processes on the molecular level to genetic, morphological, physiological and behavioral phenomena, and from their roles within encompassing ecosystem(s) to their roots in evolutionary history.

Genetics and evolution have greatly enhanced our understanding of the living. And, given the prevalence of second-gear, mechanistic thinking (sketched in §7.2.1 above), it is not surprising that both ideas have been put to use for a reductive un-

derstanding of life. Having abandoned second-gear reductionism, we may consider the thought that genetics and evolution are important aspects of life-forms, means by which life-forms come to instantiate themselves in the ways they do. There is no *need* anymore to insist that mutations are 'blind', that is, independent of the life-form in which they occur. Mutations may well happen for a reason.[386] On the other hand, there is no need to deny 'blind' mutations either. For us, what is most important is that the reasons for thinking one way or the other is not a presupposed reductionism, resting on acceptance of second-gear metaphysics, but rather a sober look at what is actually 'out there' from a non-reductionist point of view.

There are, of course, hard questions when it comes to life-forms and evolution, even after we have freed ourselves of the confusion described above, and settled for a version of essentialism that goes together quite well with evolutionary theory—an essentialism based on the notion of a third-gear life-form. Consider, again, the gradual development of instances of a new life-form out of a lineage of instances of a preexisting life-form, in accordance with orthodox evolutionary theory. We may say that as the individual organisms gradually change from members of the old life-form into members of the new life-form, the new life-form takes over the organizational work from the old life-form. Whereas earlier organisms were engaged in a cluster of life-processes that hang together in consonance with the older life-form, later organisms are engaged in a cluster of life-processes that hang together in consonance with the newer life-form. But where exactly does this 'substantial change' occur? At which point does the newer life-form take over?[387] These are hard questions indeed,

[386] In fact, there is empirical support for this way of thinking. Consider, e.g., the varieties of epigenetic inheritance—inheritance of acquired traits—that are currently being debated: Jablonka and Lamb [2005, 2007] and Jablonka and Raz [2009] present interesting findings, and argue that '[i]ncorporating epigenetic inheritance into evolutionary theory extends the scope of evolutionary thinking and leads to notions of heredity and evolution that incorporate development' [Jablonka and Raz 2009, p. 167]. They view themselves as 'challenging the modern synthesis', with its narrow focus on genes as the sole locus of evolution [see esp. Jablonka and Lamb 2008], and reintroduce ideas from the history of evolutionary theory that were banned from the modern synthesis because they do not fit the reductive project that well—e.g., Lamarckian ideas and saltational evolution (see fn. 382 on p. 303 above).

[387] The issue is reminiscent of the contemporary debate on vagueness. We are assuming that there is no metaphysical vagueness as to which life-form a given organism belongs to—that is, we are assuming an 'epistemicist' take on the issue: despite the change from earlier to later organisms being gradual, there is one precise cut-off point at which the new life-form becomes instantiated instead of the older life-form. See, e.g., Williamson [1994]. But perhaps it makes sense to consider metaphysical vagueness in the case at hand—E. Barnes and J. R. G. Williams [2009], Eklund [2013], and Lowe [2013a] for recent sympathetic discussion of such vagueness in general (and Heller [1996] for less sympathetic discussion). Notice that insofar as the issue turns on questions of vagueness, it is not peculiar to third-gear phenomena. The whole discussion on vagueness started, after all, with considerations over inanimate phenomena such as heaps (see Keefe and P. Smith [1997] for a collection of some of the historically most relevant philosophical material on vagueness, together with a series of more recent papers). More resources for thinking about speciation can be found in Strobach [2014]. Strobach introduces the notion of 'ontological retrospect',

which we cannot settle here, and that is because they are, in the end, empirical questions. There may well be no general answer to them—one has to study the life-forms involved in specific situations in order to find out, and perhaps even then is it impossible to know where the crucial point lies. But that, it is important to stress, is an epistemological problem, not a metaphysical one. It suffices for our purposes to point out that the difficulty of these questions does not warrant any skeptical conclusions with regard to life-forms. It just tells us that science is hard. Instead, skeptical conclusions are built on adherence to a form of Modern picture-based metaphysics: second-gear metaphysics.

Plasticity and multiformity are typical of the living: organisms are often very good at adapting their development and behavior to the circumstances at hand. It should not be surprising, then, that this dynamicity allows diverse life-forms to have instances that may be very similar to one another—so similar, indeed, that the existence of the one may give rise to the existence of the other (recall Marmodoro's understanding of 'instantiation by change', quoted on p. 300 above). Thus understood, evolutionary theory goes together well with the idea of life-forms—and, in this combination, it makes perfectly intelligible Darwin's quest for understanding 'the origin of species', without any need for reductionism.

7.5 Concluding Remarks

We embarked, in this chapter, on a dialectical quest to make sense of the third conceptual gear given the prevailing strong bias towards second-gear thinking. Thus, it has not been our aim to provide a positive characterization of the third conceptual gear, one that goes beyond the sketch we provided thereof in §4.4—although we have provided some further illustrations and clarifications along the way.

We have spent some of our efforts, in §7.1, trying to make clear that the third conceptual gear indeed forms a nexus of its own, which cannot be understood in second-gear terms. We did so by following Michael Thompson's discussion of a possible list of the 'signs of life', and concluded, with him, that any putative list-occupant will either not represent a true sign of life, or else will already contain the very idea of life itself.

We spent much of our efforts, in §7.2, on illustrating how third-gear phenom-

which is, roughly, the idea that there may be situations in which it is only settled later on to which species a given organism belongs (and he considers other applications of the notion, e.g., to cases involving generation and destruction, and to cases involving decisions and actions).

ena fit into a reality that also contains second-gear phenomena, resulting in two conclusions—one easy, the other hard. The first is that, of course, third-gear phenomena require second-gear objects and stuffs to be instantiated. The second is that, perhaps surprisingly, third-gear phenomena require second-gear indeterminism to be instantiated: there can be living beings at all only if the composing second-gear objects and stuffs allow to be *used as materials* by such living beings for purposes that are not yet present in those objects and stuffs taken by themselves. Along the way, we briefly looked into the advent of second-gear metaphysical dogmatism over the last centuries, which went hand in hand with an increasing preoccupation with mechanistic and reductive views on animate nature. Such dogmatism instantiates Modern picture-based metaphysical realism (as discussed in chapter 1), and we argued that it is such dogmatism which prevents a sensible view on the relation between third- and second-gear phenomena from being at all envisaged.

Furthermore, we have looked, by way of a case study, at Ruth Millikan's relationship to third-gear thought in §7.3: on the one hand, she makes good use of central notions of the third conceptual gear in her approach to questions in the philosophy of mind and language, while, on the other hand, also providing a strictly reductive theory of those third-gear notions (in 'naturalist, nonnormative, and nonmysterious terms' [Millikan 1984, p. 17], that is, in second-gear terms). We examined her reductive theory and concluded that it, in effect, founders on the same dilemma that we discussed earlier, in connection with Thompson's idea of a list of life: either Millikan's proposal as to the defining features of biological normativity do not in fact single out the realm of the living, or they do so only because third-gear concepts are being used from the start—in which case the view is not reductive anymore.

Finally, in §7.4, we devoted some attention to evolutionary theory, to clarify its relationship to the non-reductive, essentialist view on living organisms we are defending. We distinguished the way a reductive project, such as Millikan's, makes use of certain core notions of evolutionary theory (in particular, of the notion of natural selection) from the role such notions play within the theory itself: evolutionary theory is independent from second-gear reductionism. Moreover, evolutionary theory does not threaten the very idea of a life-form, as many have assumed: it is perfectly coherent with the idea of life-forms as essences (of a third-gear variety). Diversity and plasticity are hallmarks of the living, even within the lives of individual organisms, and their evolutionary origins just illustrate that fact.

Taking the third conceptual gear metaphysically seriously, which is what the Aristotelian picture suggests, implies endorsing the kind of view on reality we have

tried to circumscribe both in our general sketch of that conceptual gear in §4.4 and, from a different point of view, in the present chapter. That view on reality is anti-reductionist, non-physicalist, non-materialist. It is not dualist either, as should be clear from our discussions above: it does not place the realm of the inanimate and the realm of the animate into different worlds. It is best described as a form of conceptual realism, of hylomorphism, where the diversity of things inhabiting reality—second-gear physical objects, third-gear living beings, and presumably higher-gear entities—is mirrored in (or, rather, identical with) the diversity of conceptual gears we have to employ to understand them.

* * *

Conclusion

We have argued for an approach to metaphysics that rests on rejecting the prevailing metametaphysical orientation embodied in the Modern picture in favor of the metametaphysical orientation we recommend: the Aristotelian picture. The key *negative* insight on which that picture rests is that it is a grave mistake to assume the conceptual and the real to be, in a certain sense, disjoint—which is what the Modern picture does. For, thereby, we end up with a problematic conception of metaphysics as the project of bridging just that gap. Thus, the key *positive* insight on which the Aristotelian picture rests is that reality itself is conceptually structured: essences are concepts.

In Part I, we developed the Aristotelian picture, in its opposition to the Modern picture (chapter 1), and worked our way towards a suitable understanding both of the kind of essentialism it involves (chapter 2) and of the notion of conceptual truth it relies on (chapter 3). In Part II, we moved on to identify, on a sufficiently general level, distinct clusters of concepts, conceptual gears, that correspond to diverse, increasingly rich, and systematically interrelated ways in which concepts find expression in reality (chapter 4). We defended the resulting metaphysical picture against reductive tendencies stemming from Modern picture-based metaphysical realism, and assessed key contemporary metaphysical debates from the Aristotelian perspective we thus gained (chapters 5, 6 and 7).

This conclusion is devoted to two tasks. First, we wish to highlight the most important results of our investigations, and secondly, we wish to indicate in what ways they may be fruitfully developed further. We pursue both aims in tandem. Again, we divide our evaluation into two parts: a metametaphysical part, and a metaphysical part.

Part I: Metametaphysics

The Modern picture is characterized by the following principles, which we have extracted from key observations concerning the contemporary metaphysical debate and its main division into metaphysical realists and metaphysical anti-realists (see §1.2):

Separation	Reality as it is in itself is to be strictly distinguished from reality as it is for us.
Source	Perception consists in affectations of our senses by reality-itself, and hence constitutes our primary access to that reality.
Mind-Dependence	Our concepts are the mere products of our minds.

These principles secure that there is a gap between reality as it is in itself and reality as it is for us. Reality-for-us involves us, firstly, in that it involves *our* perceptual capacities, and secondly, in that it involves *our* conceptualizations. Thereby, it comes to stand between us and reality-itself.

Consequently, the metaphysical project becomes one of bridging the gap between reality-for-us and reality-itself—in particular, of identifying the concepts that somehow correspond to an envisaged metaphysical structure of reality-itself. Metaphysical realists understand this project to be possible: on their view, we may construct metaphysical theories that explain reality-for-us, and then, by way of an inference to the best explanation, conclude that reality-itself is as our best metaphysical theory says it is.

Both skeptics and metaphysical anti-realists understand this project to be impossible: constructing such metaphysical theories is just adding more to reality-for-us—precisely because such theorizing is inevitably *conceptual*. We cannot break out of our own minds and compare reality-for-us with reality-itself *directly*. What the metaphysical realist envisages as the metaphysical structure of reality-itself is just more conceptual structure.

The skeptic concludes that we can never know the metaphysical truth: a rather unsatisfactory result. The anti-realists, on the other hand, conclude that we were confused as to the aim of our investigations: in fact, what we aim to understand is reality-for-us, not reality-itself. This, in turn, will strike those with realist inclinations as an objectionable resignation: our aim was, after all, to understand *reality*—and we should not accept anything less.

This, then, is the menu of options to which the Modern picture gives rise (see §1.2, p. 39):

	Skepticism	*Anti-Realism*	*Realism*
Can we arrive at knowledge of reality-in-itself?	No	No	Yes
Does truth involve correspondence with reality-in-itself?	Yes	No	Yes

We may cite, here, some relevant results from Part II: there, we identified two rather generic forms which metaphysical realism can take: first-gear metaphysics and second-gear metaphysics (see below). The former assumes the metaphysical structure of reality-itself to correspond *only* to the conceptual structure of the first conceptual gear, the latter assumes the metaphysical structure of reality-itself to correspond *only* to the conceptual structure of the first and the second conceptual gears. In both cases, then, a commitment to reductionism follows: first-gear metaphysics requires reductive views about all second- and higher-gear phenomena (e.g., causality, time, life), while second-gear metaphysics requires reductive views about all third- and higher-gear phenomena (e.g., life). The basis for both views is a form of metaphysical dogmatism: a considerable restriction on the allowed conceptual resources is posited as establishing *the* standard for metaphysical theorizing.

Modern picture-based metaphysical realism distorts the metaphysical project: it encourages the formation of such metaphysical dogmas, and promotes purely speculative theorizing about reality-itself that threatens to lose all significance for our aim to understand reality. This distortion is clearly visible in the increasing amount of diverging metaphysical views one finds in the contemporary literature: there appears to be no principled way even of merely distinguishing serious metaphysical proposals from utter implausibilities (such as numberism; see §1.1.1).

For metaphysical anti-realists, on the other hand, such divergence is not in itself problematic: given that they reconceive the target of our quest for understanding reality to be not reality-itself but rather reality-for-us, choosing one or another conceptual scheme simply *makes it the case* that reality-for-us conforms to that scheme. Here, of course, we run into the problematic aspect of anti-realism: it collapses into a form of relativism that can no longer make sense of the very project of aiming to understand reality—there is no reality to *discover*, only realities to *make* (see §1.1.2).

The Aristotelian picture rejects all three defining principles of the Modern picture. Instead, it rests on the following alternatives (see §1.3):

Openness	In our experience of the world we are in direct contact with reality.
Mind-Independence	The concepts which we use to think truly about the world are part of that very world (and not mere products of our minds).
Combination	By combining the perceptual and conceptual aspects of reality in the right way, we put together what in reality is never separated.

These principles do not give rise to a gap between the conceptual and the real that then needs to be bridged by metaphysics. Instead, it takes reality itself to be concept-involving: concepts figure in the constitution of reality by being the essences of its inhabitants (in a very broad sense). Thus, when we grasp the right concepts when confronted with reality in our experience, we have succeeded in our aim to understand reality. Insofar, the Aristotelian picture captures the Modern picture-based metaphysical realists' intuition: we *can* arrive at knowledge of reality. However, contrary to what such metaphysical realists believe, there is no *further* step to be taken from our grasping the right concepts to their standing in the right sort of relation to reality-itself. Insofar, the Aristotelian picture captures the insight of the Modern picture-based metaphysical anti-realists: it is impossible to compare our concepts with a concept-less reality. However, contrary to what such anti-realists believe, that is not because we are trapped inside reality-for-us, but rather because the very idea of a concept-less reality-itself is mistaken. In McDowell's [1994a, p. 27] words: '[w]hen one thinks truly, what one thinks *is* what is the case'. What remains is the epistemically interesting but otherwise innocent fact of our fallibility: we may fail to grasp the right concepts—just as we may fail to perceive what is the case.

We have seen that the metaphysical pessimism at which Barry Stroud arrives—that the 'invulnerability' of certain basic concepts, such as causation and value, makes it impossible to assess their metaphysical credentials—results from the unfortunate combination of endorsing the Modern picture on the one hand, and acknowledging the problems to which the Modern picture-based understanding of the quest for reality gives rise on the other hand. There is no need, however, to accept such pessimism: indeed, one may take it as a reason for rejecting the Modern picture om favor of the Aristotelian picture (see §1.4).

Essentialism is a central tenet of the Aristotelian picture—it explicates the way in which concepts structure reality. Of particular importance are, in this respect, those sortal concepts that amount to ontologically fundamental sortal concepts—the con-

cepts that are elicited when asking, for a given thing, the Aristotelian question: *what is it*? They are distinguished from others in that non-fundamental sortal concepts are reducible to the fundamental ones—to belong to a non-fundamental kind of thing just is to be an instance of a certain fundamental concept that obeys additional conditions (e.g., a *bachelor* is just a *human being* that is adult, male, and unmarried). The ontologically fundamental sortal concepts are, then, those sortal concepts that are *not* thus reducible (see §2.2). More generally, whatever follows from the fact that something falls under a certain ontologically fundamental sortal concept will be essentially true of that thing. We illustrated this result by considering, in particular, a form of inference one frequently finds in the literature: the 'essentialist inference' (see §2.1). That inference draws essentialist conclusions from principles of individuation: everything has its individuating features essentially. Since the fundamental sortal concepts tell us *what* any given thing is, they embody such principles of individuation—hence it indeed follows that everything has its individuating features essentially. We noticed that we thereby arrive at a double notion of essence: the fundamental sortal concepts are general essences, their instances are individual essences—the former are characterized by a range of determinables (e.g., human beings can have *certain* hair colors, eye colors, character traits, etc.), while the latter require full determinateness (e.g., an individual human being has a *particular* hair color, eye color, etc.—see §2.4).

What is doing much of the work in this Aristotelian picture based version of essentialism is the notion of consequence, of something 'following from' something—of certain essential features 'following from' a given sortal concept. To understand that notion, and the related notion of conceptual truth, we had to remove a certain obstacle, namely, the assumption that the consequence relation is best understood on the basis of a fundamental distinction between logical form and content. Properly understood, logical form is a restriction on content, not something that is opposed to content. The logical form of a thought is that which remains if we take into consideration only certain aspects of its overall content—the logical aspects (however these are demarcated). *Logical* truths are, then, truths whose truth is guaranteed by only the logical aspects of their content (and a thought is a *logical* consequence of given premises if its truth is guaranteed by the logical aspects of its content conditional on the truth of all the premises). The fundamental notion is that of *truth guaranteed by content*, where content is now no longer understood in opposition to form, but rather as encompassing logical form. It is easily seen that the more narrow, logical notions are restrictions on the broader notions of *conceptual* truth—truth guaranteed by content as such—and *conceptual* consequence (see §3.3).

The significance of the study of logical truth and consequence, on a traditional reading of 'logical', then, is not that it captures the very phenomenon of something following from something, that is, of conceptual consequence. Rather, such a *logical* study captures the contribution to content made by certain very general and easily isolable aspects of that content—e.g., truth-functional connectives, or quantifiers. Hence it is not surprising that, as we have seen, attempts to reductively define the notions of logical truth and consequence (e.g., in model-theoretic terms, as on the Tarskian approach) fail (see §3.2). Nor is it surprising that the demarcation question for logic is so perplexing: one can restrict the aspects of content on which one wishes to concentrate in many different ways, and the result will be a 'logic' for those aspects of content. It should be clear, however, that not every such restriction lends itself to analysis in terms of familiar formal frameworks.

This understanding of conceptual truth highlights the possibility of identifying aspects of content—conceptual gears—that hang together in such a way as to be isolable for the sake of closer conceptual study (see §3.4). That is precisely what Part II was devoted to: we identified three such conceptual gears and studied their internal conceptual interdependencies and their interrelations. Thus we identified three different ways in which concepts may structure reality: in an abstract, first-gear way, in a concrete, causal-temporal second-gear way, and in a teleological, third-gear way (see chapter 4). Insofar as we conceive of logic as the conceptual study of forms based on aspects of content, then, these conceptual gears give rise to their own 'logics'—each with its own conceptual/metaphysical import. Thus, the primary focus in analytic philosophy on first-order predicate logic has done metaphysics more harm than good: it encourages a purely first-gear approach to everything (see §3.3.2).

Obviously, engaging in conceptual investigations into our three conceptual gears takes us from the metametaphysical level within which we have been working so far to the metaphysical level. Before moving on to summarize our main metaphysical conclusions, however, we wish to sketch topics for further research on the metametaphysical level. Here are four such topics, which we have not looked into even though they have surfaced during our investigations at some points:

(1) *Truth.* Which understanding of truth is the Aristotelian picture committed to? Not all truths are conceptual truths, after all, hence our construal of the notion of conceptual truth does not suffice to answer this question. It is clear that the correspondence theory as well as the currently popular 'truth-maker theory' are both wedded to the Modern picture, and hence have to be discarded. Anti-realist ('epistemic') and pragmatist theories of truth are out of the picture for similar

reasons. So, which options are left? Since we are identifying the concepts we grasp with the essences of things 'out there', it seems that what is known as the *identity theory of truth* suits our purposes.[388] The metaphysics of truth is, however, a complicated issue—the topic merits careful inquiry.[389] (We should mention that we touched upon the issue both in our introduction of the Aristotelian picture in §1.3, see fn. 65 on p. 44, and in our discussions of time, see §6.2.4, p. 243.)

(2) *Idealism*. We haven't considered the question whether the Aristotelian picture is a form of idealism or not. Given that so many radically different views have been called by that name, there is, undoubtedly, an identifiable sense of 'idealism' that applies to the Aristotelian picture. Clarification of that sense, and of its relation to other senses of 'idealism' (in particular the sense that applies to versions of Modern picture-based anti-realism), may enable a better understanding of the position of the Aristotelian picture. (We remarked on one suspicion of idealism in §1.4, p. 55.)

(3) *Epistemology of concepts*. What does it mean to 'grasp a concept', and what does it take to do so correctly? The Aristotelian picture is opposed to empiricism: concepts are not abstractions from what is given in perception. Hence grasping concepts cannot be explained by such means. Taking seriously the three Aristotelian principles given above, *Openness*, *Mind-Independence*, and *Combination*, requires that we consider the possibility that our faculties of perception and of conception are more alike than it seems. Their difference lies in the former being passive, the latter being active, but that does not mean that what conception yields is a *product* of the activity in question, such that only perception puts us into contact with reality. Rather, the activity of thinking, of grasping concepts, is one of *intuition*, to use an old-fashioned label.[390] Exploring and defending this thought would greatly clarify the implications of the Aristotelian picture outside of the metaphysical sphere we have been concerned with, and connect it with issues in epistemology and the philosophy of mind. (We touched upon the issue in our introduction of the distinction between concepts and conceptions in §1.3, p. 44, and in our sketch of Millikan's view on concepts (or 'unicepts') in §7.3.1.)

[388] See, e.g., Baldwin [1991], Hornsby [1997], Candlish [1999], Engel [2001], McDowell [1994a, esp. lecture 1, 2005], and F. Ellis [2005, esp. ch. 5].

[389] We have ignored deflationism: it is either an unfortunate way of presenting the identity theory we recommend, or nothing but a semantic observation concerning the word 'true' without much metaphysical significance. However, we cannot go into the matter here, obviously.

[390] We should emphasize that we here do *not* have in mind the construal of intuitions as 'intellectual seemings' promoted by, e.g., Bealer [1996, 1998, 1999]. We are talking about intuition as a capacity to grasp concepts, not entire propositions.

(4) *Logic*. We have defended an understanding of logic which we were only able to present in briefest outline. It would be helpful to see how that understanding of logic may shed light on issues in philosophical logic—most importantly, of course, on the demarcation question, but also on the relation between model theory and proof theory and on the status of logic as a normative or a descriptive discipline. (We have touched upon the relevant issues in chapter 3, see especially §3.2.3 and §3.3. MacFarlane's [2000, 2009] contributions to the issue are especially valuable in this respect.)

As one looks at this list of topics, it becomes apparent that the Aristotelian picture is not just a metametaphysical picture, but rather a general *philosophical* picture—it is just that we have been developing and explicating it from the standpoint of a metaphysician. (And, with Aristotle, we could add that metaphysics is, after all, first philosophy.[391])

Part II: Metaphysics

The Aristotelian picture encourages the search for conceptual distinctions and interrelations: insofar as we are grasping the right concepts, such conceptual work simply *is* metaphysical work. In particular, as metaphysicians, we are interested not so much in exactly which kinds of things we find 'out there', in all their diversity and specificity, but rather in the most general conceptual divisions one can make. These are often implicit in our understanding of reality, both in ordinary life and in the sciences.

On a very abstract level, we find the pure categories of an *entity*, that of which one can predicate something, and of a *feature* or *relation*, that which one can predicate of entities. Given the essentialism we established in Part I, we also take into account the fundamental sortals under which such entities fall: the *type*. Two forms of predication become salient: *sortal predication*, by which one says of a thing *what it is*, and *feature predication*, that by which one says of a thing that it exemplifies a certain feature. Both forms of predication are, on this abstract level, atemporal.

We have called this general conceptual structure the *first conceptual gear* (see §4.2), and we observed that all of its elements occur, in a transformed and enriched way, in another general conceptual structure, the *second conceptual gear* (see §4.3). Here, we find *physical objects* that persist through time: their existence is no longer atemporal

[391] See, for Aristotle's characterization of the 'science' of metaphysics, *Metaphysics* [1998, p. 79, IV.1, 1003a22–32].

but rather temporally bound. The sortal concepts under which they fall may be named *natural kinds*. Their features now differentiate into states and activities; the former are ascribed by use of a tensed form of predication, the latter are ascribed by use of a tensed and aspected form of predication. Physical objects are capable of engaging in certain activities: they have *causal powers* which they manifest when the right circumstances obtain—and the manifestations, the activities themselves, are *processes*, in themselves entities of which one can predicate.

A further transformation of this already much richer conceptual structure is what we find in the *third conceptual gear* (see §4.4): here, we find *organisms*, whose persistence amounts to living. The fundamental sortal concepts under which such organisms fall are *life-forms*. Organisms are engaged in *life-processes*: processes that play a certain role for the life of the organism in question and are therefore characterized by a certain normativity: organisms *should* perform the life-processes that their life-forms dictate. Life-forms thus embody a remarkably versatile organizational principle that expresses itself in the way their instances grow, develop, organize and maintain themselves, and reproduce. The normativity in question characterizes a further differentiation in the form of predication involved, beyond tense and aspect.

The following scheme gives an overview of these three conceptual gears:

	First gear	**Second gear**	**Third gear**
form of predication:	tenseless	tensed/aspected	normative
applies to:	entity	object	organism
predicates involved:	feature	state/process	life-process
sortals involved:	type	natural kind	life-form
mode of being:	to exist	to persist	to live
engaged in:	—	process	life-process
characteristic:	geometry	causality/temporality	teleology/normativity
	Humeanism	*anti-Humeanism*	

We find that we can sensibly make these conceptual differentiations, and we find that we understand reality by using all three conceptual gears (and presumably others as well). On the Aristotelian picture, that means that we should take these conceptual structures seriously—metaphysically seriously, if you like.

On the other hand, the distinction between these three conceptual gears aids our understanding of what is at stake in many of the contemporary debates in metaphysics. As we said, there are Modern picture-based metaphysical realisms based on restrictions to the first or the second conceptual gear only—resulting in the need for reductive theories of higher-gear phenomena.

Given its skeptical attitude towards 'necessary connections' and the like, a broadly Humean take on metaphysics is attractive for first-gear metaphysicians: only first-gear concepts are acceptable by their lights, and Humeanism allows for at most such concepts in its view on reality-itself. However, the fit is not perfect: once we discover that it is conceptual connections in general to which the Humean is opposed, we find that there is considerable pressure for him to abandon realist metaphysics altogether, in favor of anti-realism (see §5.2.4). Still, those restricting themselves to first-gear thought only, such as David Lewis, may reasonably be understood to endorse a form of Humeanism—as, indeed, they usually characterize themselves (viz., Lewis's grand thesis of 'Humean supervenience' [see 1986b, pp. ix–x]).

On such a first-gear view, all of the typically second-gear concepts receive a reductive treatment. A straightforward implementation of this view embodies an understanding of time that includes eternalism, B-theory, and perdurantism; and a non-governing, regularities-based conception of laws of nature and causality (see §5.1 and 5.2, respectively). The result is that temporal truths are analyzed in atemporal terms (in accordance with the first-gear form of predication), and that the distinction between objects and processes collapses (in accordance with the single, first-gear notion of an entity). The entire causal-temporal order of the second conceptual gear is reduced to a geometrical array of independent matters of fact.

What the first-gear metaphysician offers is a simulation of second-gear concepts in a first-gear setting. And he can do so with impressive sophistication, as careful study of elaborate first-gear metaphysical views shows. In fact, the simulation can be so sophisticated, that even its proponents are not aware of the fact that it *is* a simulation. One may defend what appear to be the opposites of typical first-gear views—presentism, A-theory, endurantism, primitive necessary connections, and the like—without escaping from the realm of first-gear thought. David Armstrong's necessitarian understanding of laws of nature and causation is a case in point (see §6.1), as is the tense-logical understanding of A-theory defended by Arthur Prior, Kit Fine and others (at least, on one reasonable understanding of their views—see §6.2.4).

We thus find that the contemporary controversies between eternalism and presentism, between A-theory and B-theory, between perdurantism and endurantism, and between Armstrongian necessitarians and Humeans are often strikingly confused: whereas what motivates all of these controversies is the question whether or not one should take the second conceptual gear metaphysically seriously, the contestants often turn out to be battling within a purely first-gear conceptual arena (see §6.3).

Simultaneously, we learn that the distinction between the first and the second conceptual gears is indeed a very fundamental one. Although there are many different ways of simulating second-gear concepts within the first conceptual gear, there is not much room for different views *within* the second conceptual gear. Presentism, A-theory and endurantism are mandatory—though one should take care not to misconstrue either of these aspects of second-gear thought in first-gear terms. We may prevent such misconstruals by putting the second-gear form of predication on center stage: it connects physical objects with their states, powers, and processes in a tensed way, without there being any underlying tenseless way of unifying the same ingredients. Obviously, B-theory contradicts this second-gear starting point right away, with its dismissal of tense. Likewise, eternalism is only an option if one understands temporal truths as, in the end, atemporal truths located at one or another point on a time-line—a paradigmatically first-gear thought. Therefore, there is nothing to contrast presentism with on the second conceptual gear (see §6.2.3). And with eternalism goes perdurantism: it rests on a fundamental notion of atemporally occupying temporal locations, which is unavailable on the second conceptual gear (see §6.2.5). That is why we should replace these three classical oppositions—between A-theory and B-theory, eternalism and presentism, and perdurantism and endurantism—with the one opposition that matters: whether or not we should take the second conceptual gear metaphysically seriously.

On the topic of causality we draw similar conclusions (see §6.1.5). As long as one sticks to the thought that causation is a two-place relation on a given domain of events (or whatever the relata are thought to be), one is trapped within first-gear thought. The second-gear concept of causation is that of causal production: by manifesting their causal powers, presently existing physical objects produce and undergo changes. This second-gear understanding of causality obviously depends on the presentist, A-theoretic, endurantist understanding of time we sketched. And the dependence is mutual: an allegedly temporal reality consisting of a series of static arrangements of states, without any causal processes, collapses into an atemporal series of abstract, first-gear realities. Time and change are two sides of one and the same second-gear coin—as is mirrored in the second-gear form of predication, which displays both tense (time) and aspect (change).

Despite the broad acceptance of a Humean, first-gear understanding of metaphysics amongst contemporary analytic philosophers, heritage of the empiricist roots of analytic philosophy, there is a respectable (and growing) anti-Humean, second-gear camp—witness, in particular, the growing number of defenders of a powers-

based metaphysics. When moving on to the shift from second- to third-gear thought, however, the situation is different: defenders of a truly anti-reductionist (i.e., third-gear) view on life are much rarer than defenders of a truly anti-reductionist (i.e., second-gear) view on causation and time.

Accordingly, we argued for the autonomy of the third conceptual gear largely by arguing against the prevailing second-gear metaphysical dogmatism, clarifying what such autonomy amounts to on the way. As with first-gear simulations of second-gear concepts, second-gear simulations of third-gear concepts can be easily distinguished from truly third-gear concepts: the former fail to display the characteristic features of the latter. In this case, those features are the normativity and teleology that characterize the realm of the living. Thus, if the second-gear thinker tries to identify life by, e.g., pointing towards the impressive sort of organization that living things display, such a statement turns out either to fail in identifying life, or succeeds trivially by tacitly relying on the very idea of life (Thompson's 'sub-metaphysical Scylla' and 'tautological Charybdis'—see §7.1). In the former case, the organization in question may just as well be displayed by the non-living. In the latter case it simply highlights a way in which life manifests itself—by organizing its materials in a peculiar way. We observed that Ruth Millikan's carefully constructed reductive account of third-gear normativity in second-gear terms indeed illustrates this result: her definition of direct proper function, which is where the normativity comes in, turns out to find application wherever contingent stability of a suitable kind is found—which may clearly happen outside of the realm of the living (see §7.3).

Organisms, instantiations of life-forms, indeed organize the matter they are made of in accordance with their life-form. Thereby, such matter no longer *simply* obeys the second-gear laws it falls under: it now *also* obeys the third-gear principle of organization, the life-form, into which it is taken up (see §7.2.2). Thus, third-gear life-forms can only instantiate themselves if there is room for them to organize the underlying matter in this additional way. Such room is only present if the second-gear laws in question are not deterministic, that is, if the future is open (or, to be more precise, if there are spontaneous or indecisive causal powers). To use the analogy we adopted from Elizabeth Anscombe (see §7.2.1, p. 267): the second-gear laws have to be like the rules of chess, so that organisms can organize their underlying matter, make their 'moves', within the space of real possibilities that these laws ground (see also §6.2.6). Thereby, everything that an organism does, from the internal processes on the molecular microlevel to the interactions with its environment on the ecological macrolevel, is not merely a second-gear causal process devoid of teleology and

normativity but rather a third-gear life-process, which *aims* to and *ought* to unfold in such a way as to further its hosting organism.

To understand what this means, we need to free ourselves from the strong, second-gear metaphysical dogma which has been ingrained in our thinking ever since the rise of the Modern picture in the early modern era (see §7.2.1). That dogma dictates that we be able to understand life in mechanistic (second-gear) terms. Interestingly, a proper understanding of that very requirement of mechanistic explanation already provides the materials we need to make sense of our strong anti-reductionist claim. For, in the end, the second-gear laws of nature are *conceptual truths* concerning second-gear physical objects, the stimulus conditions for their causal powers, and the processes in which the manifestations of such powers consist. Such conceptual truths govern what happens: they form part of the essences of the objects, powers, and processes in question, and as such act as *formal causes* determining what can and what cannot happen (see §4.3, p. 139, and also §6.1.4, p. 199). The way third-gear life-forms operate is, viewed from this suitably general angle, precisely analogous: they, too, act as formal causes determining what can and what cannot happen (see §7.2.2, p. 270). The difference is, of course, that third-gear life-forms require underlying second-gear materials, whereas this is not the case for those second-gear materials themselves (see §7.4.2, p. 306). That is why there may seem to be a tension between the second-gear laws and the operations of third-gear life-forms. That tension disappears once we see that on an indeterministic picture, there is room for such life-forms to operate on *without* breaking the underlying, second-gear laws. Failure to understand the formal aspect of causation, and the way it makes intelligible the structuring of reality by concepts both on the second and on the third conceptual gears, leads to caricatures of views like vitalism as involving hidden, ephemeral forces or stuffs which are still thought of in second-gear terms.

There is disagreement amongst biologists and philosophers of biology over the question what biological species are—but, at least, there appears to be a widespread consensus on one thesis: that biological essentialism is wrong. This thesis is often based on evolutionary theory: since species evolve, they cannot be static, hence essentialism fails (see §7.4.2). This is a fallacy—precisely in the realm of biology, where we find life-forms exhibiting remarkable degrees of versatility and dynamicity, it seems possible that instances of a given life-form may, over time, due to environmental pressures and the like, give rise to instances of a different life-form. Evolution, however gradually it proceeds, does not contradict essentialism. There is, then, no 'species problem' in this skeptical sense. There is, of course, a serious question as to

how we may identify and distinguish life-forms: an epistemological problem. Since life-forms organize themselves on many levels at once—from the molecular up to the ecological—one needs to get to know the principle of organization involved on all of these levels in order to begin tackling this question. Life-forms are very complex and rich concepts containing a diversity of teleologically interrelated factors. They should not be approached with an all too simple account such as 'genetic essentialism' in mind. In short: science is hard, but that is no reason to abandon essentialism.

There appears to be some agreement on another thesis: the non-reducibility of biology to chemistry and physics. That thesis is, however, usually put in a puzzlingly vague manner, largely because its proponents normally point to mere epistemological or methodological non-reducibility, sometimes coupled with a few statements about 'downward causation' and 'emergence' that are never satisfactorily spelled out (see §7.2.2, p. 271). This betrays an underlying failure to get rid of the materalistic-mechanistic metaphysical dogma that stands in the way of a satisfactory understanding of life. The Aristotelian picture rejects such a dogma and takes the third conceptual gear at face value: thereby it provides a straightforward content for the claim the biological anti-reductionists appear to be after. The structure, then, is rather like what we found in the debates between Humeans and anti-Humeans: there, we concluded that what we have is a dispute over the question whether or not we should take the second conceptual gear metaphysically seriously. The issue gets much more complicated than necessary because this bottom line of the various disputes is not understood properly. Similarly, the issue over the reducibility of biological phenomena is a matter of taking (or not taking) the third conceptual gear metaphysically seriously. And, again, the issue becomes much more complicated than necessary because this bottom line of the dispute is not understood properly (see §7.5).

On the Aristotelian picture, such confusions do not arise. There is no question as to whether or not we should take one or another conceptual gear 'metaphysically seriously', because applying them *already is* taking them seriously. There is, of course, room for skeptical questions and grand revisions of the ways we think about reality—but not based on a Modern picture separation between reality-for-us and reality-itself. If we learn that certain of our conceptions fail to connect to concepts that indeed structure reality—as, e.g., happened when the belief in witchcraft was gradually abandoned—we simply *stop* using them (see §1.4, p. 48). There is no sign of such elimination in our understanding of the realm of the living, despite the strong reductive tendencies spawned by the rise of molecular biology (see §7.2.1 and §7.4.2).

There are many questions to be asked concerning our system of conceptual gears. We highlight the most important of them to indicate in which way our views may be most fruitfully developed further:

(1) *Further gears*. It looks like there are further conceptual gears: one for sentient life, and one for rational life. Consciousness and self-consciousness appear to be the corresponding forms of being (analogous to persistence and living on the second and third gears). Exploration of these conceptual gears, and of the light they may shed on contemporary debates in the relevant areas of philosophy, is of importance for a better grasp of the Aristotelian picture. (Thompson explores a conception of ethics as grounded in a specific conceptual gear for rational life—see, e.g., M. Thompson [2008])

(2) *Abstract objects*. We noticed that the first conceptual gear applies to the abstract entities of mathematics. This observation merits further exploration, to clarify what place such objects have within the Aristotelian picture—in particular, on whether they should be understood to be purely conceptual (as our conjecture in §4.2, p. 130 suggests). Such an exploration would clarify the implications of the Aristotelian picture for the foundations of mathematics.

(3) *Second-gear objects and stuffs*. We have not discussed second-gear physical objects that much—our toy example was that of a ball, and we remarked that it is inadequate because it is an artifact. On reflection, it turns out to be rather difficult to come up with good examples of second-gear objects. Perhaps the particles of physics will do, or stars, galaxies, and the like. By contrast, examples of *stuffs* are much easier: water, gold, helium, magma, granite, and the like—yet we left those out of our picture to forgo unnecessary complications. The question is, then, what a good view on second-gear physical objects and their relations to such stuffs looks like.

(4) *Causation*. We have sketched a view on causation as involving an effective and a formal factor: the effective cause brings about the conditions under which powers are manifested, processes initiated, according to the conceptual truths the relevant essences embody (the formal cause). This view may be spelled out in much greater detail, and, in particular, it should be compared to the actual practice of the sciences—for, if it is correct, scientific research is in large part about finding out which are the stimulus conditions for which manifestations, which objects and stuffs have which powers and further properties, etc.

(5) *Life-forms*. The concept of a life-form is a pure third-gear concept. It is important to clarify its content by investigating into the actual classifications of biologists—do

life-forms encompass much more than one species, or is there, roughly, a one-to-one correspondence? And, assuming that a fourth conceptual gear pertaining to conscious beings endowed with perception and behavior has been developed (see above), how does this mesh with the fact that the plant/animal distinction in biology has long been replaced by allegedly more fundamental, empirically supported distinctions such as that between prokaryotes and eukaryotes? In short, a lot of clarificatory work may be done in order to connect the concept of life-forms with our biological knowledge.

This list of possibly fruitful avenues of further research again highlight the fact that the Aristotelian picture is not just a metametaphysical picture, but rather a general philosophical picture—one that can tie together many different branches of philosophy.

To Conclude

Throughout our discussions, we have aimed at outlining the Aristotelian picture and the view of reality to which that picture gives rise. The Aristotelian picture portrays reality as at least as complex, multifaceted, and rich as the conceptual resources we employ in our quest to understand it. It is opposed to the Modern picture, which downgrades this complex, multifaceted and rich reality to the status of a mere reality-for-us, and then asks the skeptical question what, if anything, corresponds to an imagined, non-conceptual reality-itself. On the Aristotelian picture there is no such gap to be bridged. Thus, there turns out to be no reason at all to accept the abstract, almost mathematical view on reality of which first-gear metaphysics dreams, or the purely mechanical view on reality which second-gear metaphysicians endorse, or any other metaphysics based on unmotivated conceptual restrictions. The Aristotelian picture takes our quest to understand reality at face value, and thereby reveals what is already open to view: that reality is a place in which we may find ourselves to be at home.

* * *

References

Adams, Robert M.

1986 "Time and Thisness," *Midwest Studies in Philosophy*, 9, pp. 315–29. (Cited on p. 243.)

Albert, David Z.

2001 *Time and Chance*, Harvard University Press, Cambridge, Mass. (Cited on pp. 172, 173.)

Albert, David Z. and Barry Loewer

1988 "Interpreting the Many-Worlds Interpretation," *Synthese*, 77, pp. 195–213. (Cited on p. 253.)

Anscombe, G. E. M.

1971 "Causality and Determinism," in Anscombe [1981], pp. 133–47. (Cited on pp. 195, 267, 277, 381.)

1981 *Collected Philosophical Papers, Volume 2: Metaphysics and the Philosophy of Mind*, Wiley-Blackwell, Oxford. (Cited on p. 327.)

Aristotle

1963 *Categories and De Interpretatione*, Translated by John L. Ackrill, Oxford University Press, Oxford. (Cited on pp. 18, 177, 281.)

1984 *On the Soul*, Translated by John A. Smith, Princeton University Press, Princeton. (Cited on pp. 18, 46, 151.)

1996 *Physics*, Translated by Robin Waterfield, Oxford University Press, Oxford. (Cited on pp. 18, 139.)

1998 *Metaphysics*, Translated by Hugh Lawson-Tancred., Penguin Books, London. (Cited on pp. 1, 2, 18, 82, 139, 208, 318, 355.)

2000 *Nicomachean Ethics*, Translated by Roger Crisp, Cambridge University Press, Cambridge. (Cited on p. 113.)

Armstrong, David M.

1961 *Perception and the Physical World*, Routledge & Kegan Paul, London. (Cited on p. 192.)

1962 *Bodily Sensations*, Routledge & Kegan Paul, London. (Cited on p. 192.)

1973 *Belief, Truth and Knowledge*, Cambridge University Press, London. (Cited on p. 192.)

1978a *Universals and Scientific Realism, Volume 1: Nominalism and Realism*, Cambridge University Press, Cambridge. (Cited on p. 23.)

1978b *Universals and Scientific Realism, Volume 2: A Theory of Universals*, Cambridge University Press, Cambridge. (Cited on pp. 23, 189, 196.)

1983 *What is a Law of Nature?* Cambridge University Press, Cambridge. (Cited on pp. 23, 127, 189, 196, 200.)

1993 "The Identification Problem and the Inference Problem," *Philosophy and Phenomenological Research*, 53(2), pp. 421–2. (Cited on pp. 190, 192.)

1997 *A World of States of Affairs*, Cambridge University Press, Cambridge. (Cited on pp. 23, 127, 184, 189, 191, 192, 196, 200, 201.)

2004a "Going Throug the Open Door Again: Counterfactual Versus Singularist Theories of Causation," in Collins *et al.* [2004], ch. 19, pp. 445–57. (Cited on pp. 193, 194.)

2004b *Truth and Truthmakers*, Cambridge University Press, Cambridge. (Cited on pp. 23, 38, 192.)

Armstrong, David M.
2010 *Sketch for a Systematic Metaphysics*, Oxford University Press, New York. (Cited on pp. 23, 127, 192.)

Asay, Jamin and Sam Baron
2014 "The Hard Road to Presentism," *Pacific Philosophical Quarterly*, 95(1), (in press). (Cited on p. 226.)

Atran, Scott
1998 "Folk Biology and the Anthropology of Science: Cognitive Universals and Cultural Particulars," *Behavioral and Brain Sciences*, 21(04), pp. 547–69. (Cited on p. 117.)

Ayala, Francisco José and Robert Arp
2010 *Contemporary Debates in Philosophy of Biology*, Blackwell Publishing, Malden, Mass. (Cited on pp. 299, 330, 333, 339, 349.)

Ayer, Alfred J.
1956a *The Problem of Knowledge*, Penguin Books, London. (Cited on p. 215.)
1956b "What is a Law of Nature?" *Revue Internationale de Philosophie*, 10, pp. 144–65. (Cited on p. 178.)
1979 "Reply to Wiggins," in MacDonald [1979]. (Cited on p. 64.)

Bailey, Andrew M.
2010 "No Bare Particulars," *Philosophical Studies*, 158(1), pp. 31–41. (Cited on p. 79.)

Baldwin, Thomas
1991 "The Identity Theory of Truth," *Mind*, 100(1), pp. 35–52. (Cited on p. 317.)
1999 "Back to the Present," *Philosophy*, 74(2), pp. 177–97. (Cited on p. 214.)

Barnes, Barry and John Dupré
2008 *Genomes and What to Make of Them*, University of Chicago Press, Chicago. (Cited on p. 266.)

Barnes, Elizabeth and Ross P. Cameron
2009 "The Open Future: Bivalence, Determinism and Ontology," *Philosophical Studies*, 146(2), pp. 291–309. (Cited on p. 249.)
2011 "Back to the Open Future," *Philosophical Perspectives*, 25(1), pp. 1–26. (Cited on p. 249.)

Barnes, Elizabeth and J. Robert G. Williams
2009 "Vague Parts and Vague Identity," *Pacific Philosophical Quarterly*, 90(2), pp. 176–87. (Cited on p. 307.)

Baron, Sam
2012 "Presentism and Causation Revisited," *Philosophical Papers*, 41(1), pp. 1–21. (Cited on p. 226.)

Bateson, William
1894 *Materials for the Study of Variation*, Macmillan, London. (Cited on p. 303.)

Bealer, George
1996 "On the Possibility of Philosophical Knowledge," *Philosophical Perspectives*, 10, pp. 1–34. (Cited on p. 317.)
1998 "Intuition and the Autonomy of Philosophy," in DePaul *et al.* [1998], ch. 12, pp. 201–40. (Cited on p. 317.)
1999 "A Theory of the A Priori," *Philosophical Perspectives*, 13, pp. 29–56. (Cited on p. 317.)

Bechtel, William
2013 "Addressing the Vitalist's Challenge to Mechanistic Science: Dynamic Mechanistic Explanation," in Normandin *et al.* [2013], ch. 14, pp. 354–70. (Cited on p. 268.)

Bedau, Mark A.
1997 "Weak Emergence," *Philosophical Perspectives*, 11, pp. 375–99. (Cited on p. 271.)
2002 "Downward Causation and the Autonomy of Weak Emergence," *Principia*, 6(1), pp. 5–50. (Cited on p. 271.)

Beebee, Helen
2000 "The Non-Governing Conception of Laws of Nature," *Philosophy and Phenomenological Research*, 61(3), pp. 571–94. (Cited on p. 172.)

Beebee, Helen and Julian Dodd
2005 (eds.), *Truthmakers: The Contemporary Debate*, Oxford University Press, New York. (Cited on p. 38.)

Beiser, Frederick C.
2008 (ed.), *The Cambridge Companion to Hegel and Nineteenth-Century Philosophy*, Cambridge University Press, Cambridge. (Cited on p. 339.)

Belnap, Nuel
1992 "Branching Space-Time," *Synthese*, 92(3), pp. 385–434. (Cited on p. 253.)
2012 "Newtonian Determinism to Branching Space-Times Indeterminism in Two Moves," *Synthese*, 188(1), pp. 5–21. (Cited on p. 253.)

Benovsky, Jiri
2009a "On (Not) Being in Two Places at the Same Time: An Argument Against Endurantism," *American Philosophical Quarterly*, 46(3), pp. 239–48. (Cited on p. 206.)
2009b "Presentism and Persistence," *Pacific Philosophical Quarterly*, 90(3), pp. 291–309. (Cited on p. 164.)

Beurton, Peter J., Raphael Falk, and Hans-Jörg Rheinberger
2000 (eds.), *The Concept of the Gene in Development and Evolution: Historical and Epistemological Perspectives*, Cambridge University Press, Cambridge. (Cited on p. 336.)

Bickhard, Mark H.
2011a "Does Process Matter? An Introduction to the Special Issue on Interactivism," *Axiomathes*, 21, pp. 1–2. (Cited on p. 137.)
2011b "Some Consequences (and Enablings) of Process Metaphysics," *Axiomathes*, 21, pp. 3–32. (Cited on p. 137.)

Bird, Alexander
1998 "Dispositions and Antidotes," *The Philosophical Quarterly*, 48(191), pp. 227–34. (Cited on p. 197.)
2005 "Laws and Essences," *Ratio*, 18(4), pp. 437–61. (Cited on pp. 127, 198.)
2007 *Nature's Metaphysics: Laws and Properties*, Oxford University Press, New York. (Cited on pp. 23, 72, 127, 198.)

Blackburn, Simon
1993 *Essays in Quasi-Realism*, Oxford University Press, New York. (Cited on p. 32.)

Blackson, Thomas A.
1992 "The Stuff of Conventionalism," *Philosophical Studies*, 68(1), pp. 65–81. (Cited on p. 31.)

Boghossian, Paul A.
1989 "The Rule-Following Considerations," *Mind*, 98(392), pp. 507–49. (Cited on p. 283.)

Bourne, Craig
2006 *A Future for Presentism*, Clarendon Press, Oxford. (Cited on pp. 226, 227.)

Braithwaite, Richard B.
1927 ""The Idea of Necessary Connexion" (I)," *Mind*, 36(144), pp. 467–77. (Cited on p. 171.)
1928 ""The Idea of Necessary Connexion" (II)," *Mind*, 37(145), pp. 62–72. (Cited on p. 171.)

Broad, Charles D.
1923 *Scientific Thought*, Routledge & Kegan Paul, London. (Cited on pp. 220, 221, 243.)
1925 *The Mind and Its Place in Nature*, Routledge & Kegan Paul, London. (Cited on p. 271.)
1938 *Examination of McTaggart's Philosophy*, Cambridge University Press, Cambridge. (Cited on p. 214.)

Brown, Richard
2014 (ed.), *Consciousness Inside and Out: Phenomenology, Neuroscience, and the Nature of Experience*, Springer, Dordrecht. (Cited on p. 344.)

Burian, Richard M.
2010 "Selection Does Not Operate Primarily on Genes," in Ayala *et al.* [2010], ch. 8, pp. 141–64. (Cited on p. 298.)
Burke, Michael B.
1994 "Dion and Theon: An Essentialist Solution to an Ancient Puzzle," *The Journal of Philosophy*, 91(3), pp. 129–39. (Cited on p. 24.)
Butler, Joseph
2009/1827 *Fifteen Sermons Preached at the Rolls Chapel*, Dodo Press. (Cited on p. 75.)
Callender, Craig
2002 *Time, Reality & Experience*, Cambridge University Press, New York. (Cited on pp. 349, 350.)
2011a "Philosophy of Science and Metaphysics," in French *et al.* [2011], ch. 3, pp. 33–54. (Cited on pp. 18, 27.)
2011b (ed.), *The Oxford handbook of philosophy of time*, Oxford University Press, Oxford. (Cited on p. 354.)
Callender, Craig and Jonathan Cohen
2010 "Special Sciences, Conspiracy and the Better Best System Account of Lawhood," *Erkenntnis*, 73(3), pp. 427–47. (Cited on pp. 172, 174.)
Cameron, Ross P.
2012 "Why Lewis's Analysis of Modality Succeeds in its Reductive Ambitions," *Philosopher's Imprint*, 12(8). (Cited on p. 20.)
Campbell, Keith
1990 *Abstract Particulars*, Blackwell Publishing, Oxford. (Cited on p. 70.)
Candlish, Stewart
1999 "A Prolegomenon to an Identity Theory of Truth," *Philosophy*, 74(288), pp. 199–220. (Cited on p. 317.)
Caplan, Ben and David Sanson
2011 "Presentism and Truthmaking," *Philosophy Compass*, 6(3), pp. 196–208. (Cited on p. 226.)
Carnap, Rudolph
1946 "Modalities and Quantification," *The Journal of Symbolic Logic*, 11, pp. 33–64. (Cited on p. 106.)
1950 "Empiricism, Semantics and Ontology," *Revue Internationale de Philosophie*. (Cited on pp. 29, 106.)
1956/1947 *Meaning and Necessity: A Study in Semantics and Modal Logic*, 2nd edition, Chicago University Press. (Cited on p. 106.)
Cartwright, Nancy
1984 *How the Laws of Physics Lie*, Oxford University Press, Oxford. (Cited on p. 169.)
Chakravartty, Anjan
2007 *A Metaphysics for Scientific Realism: Knowing the Unobservable*, Cambridge University Press. (Cited on p. 198.)
Chalmers, David J.
2012 *Constructing the World*, Oxford University Press, New York. (Cited on p. 127.)
Chalmers, David J., David Manley, and Ryan Wasserman
2009 *Metametaphysics: New Essays on the Foundations of Ontology*, Oxford University Press, New York. (Cited on pp. 4, 18, 337, 350.)
Charles, David
2002 *Aristotle on Meaning and Essence*, Oxford University Press, New York. (Cited on p. 302.)
Chisholm, Roderick M.
1990a "Events Without Times: an Essay on Ontology," *Noûs*, 24(3), pp. 413–27. (Cited on p. 216.)

Chisholm, Roderick M.
1990b "Referring to Things That No Longer Exist," *Philosophical Perspectives*, 4, pp. 545–56. (Cited on p. 216.)
Choi, Sungho
2003 "Improving Bird's Antidotes," *Australasian Journal of Philosophy*, 81(4), pp. 573–80. (Cited on p. 197.)
Chouard, Tanguy
2010 "Evolution: Revenge of the Hopeful Monster," *Nature News*, 463(7283), pp. 864–7. (Cited on p. 303.)
Ciuni, Roberto, Kristie Miller, and Giuliano Torrengo
2013 (eds.), *New Papers on the Present*, Philosophia Verlag, Munich. (Cited on pp. 341, 354.)
Claridge, Michael F., H. A. Dawah, and M. R. Wilson
1997 (eds.), *Species: The Units of Biodiversity*, Chapman & Hall, London. (Cited on p. 342.)
Clayton, Philip and Paul C. W. Davies
2006 (eds.), *The Re-Emergence of Emergence: the Emergentist Hypothesis From Science to Religion*, Oxford University Press, New York. (Cited on p. 332.)
Cohen, Jonathan and Craig Callender
2009 "A Better Best Systems Account of Lawhood," *Philosophical Studies*, 145(1), pp. 1–34. (Cited on pp. 172–174, 178.)
Cohen, Robert S., Paul K. Feyerabend, and Marx W. Wartofsky
1976 (eds.), *Essays in Memory of Imre Lakatos*, D. Reidel, Dordrecht. (Cited on p. 347.)
Collins, John, Ned Hall, and L. A. Paul
2004 (eds.), *Causation and Counterfactuals*, MIT Press, Cambridge, Mass. (Cited on pp. 186, 327, 340.)
Colodny, Robert
1962 (ed.), *Frontiers of Science and Philosophy*, University of Pittsburgh Press, Pittsburgh. (Cited on p. 350.)
Correia, Fabrice and Andrea Iacona
2013 *Around the Tree: Semantic and Metaphysical Issues Concerning Branching and the Open Future*, Springer Verlag, Dordrecht. (Cited on p. 342.)
Correia, Fabrice and Sven Rosenkranz
2012 "Eternal Facts in an Ageing Universe," *Australasian Journal of Philosophy*, 90(2), pp. 307–20. (Cited on pp. 207, 239, 240.)
Craig, William L.
2000 *The Tensed Theory of Time*, Kluwer, Dordrecht. (Cited on p. 220.)
Crick, Francis and James D. Watson
1953a "Genetical Implications of the Structure of Deoxyribonucleic Acid," *Nature*, 171, pp. 964–7. (Cited on p. 266.)
1953b "Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid," *Nature*, 171, pp. 737–8. (Cited on p. 266.)
Crisp, Thomas M.
2003 "Presentism," in Loux *et al.* [2003], ch. 8, pp. 211–45. (Cited on p. 208.)
2004 "On Presentism and Triviality," in Zimmerman [2004], ch. 2, pp. 15–20. (Cited on p. 216.)
Curtis, Helena
1979 *Biology*, Worth, New York. (Cited on p. 260.)
Darwin, Charles
1859 *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*, John Murray, London. (Cited on pp. 301, 303.)

Davidson, Donald

1967 "Causal Relations," *The Journal of Philosophy*, LXIV(21), pp. 691–703. (Cited on p. 184.)
1973 "In Defence of Convention T," in Davidson [1984], ch. 5, pp. 65–75. (Cited on p. 100.)
1984 *Inquiries Into Truth and Interpretation*, Oxford University Press, Oxford. (Cited on p. 332.)
1987 "Knowing One's Own Mind," *Proceedings and Addresses of the American Philosophical Association*, 60(3), pp. 441–58. (Cited on p. 291.)
2006 *The Essential Davidson*, Oxford University Press. (Cited on p. 347.)

Davies, Paul C. W.

1999 *The Fifth Miracle: The Search for the Origin and Meaning of Life*, Simon & Schuster, New York. (Cited on p. 261.)
2006 "The Physics of Downward Causation," in Clayton *et al.* [2006], ch. 2, pp. 35–52. (Cited on p. 261.)

Davis, Richard B. and David S. Brown

2007 "A Puzzle for Particulars?" *Axiomathes*, 18(1), pp. 49–65. (Cited on p. 79.)

Dawkins, Richard

1996 *The Blind Watchmaker: Why the Evidence of Evolution Reveals a Universe Without Design*, Norton, New York. (Cited on pp. 294, 297.)

De Clerq, Rafael

2006 "Presentism and the Problem of Cross-Time Relations," *Philosophy and Phenomenological Research*, 72(2), pp. 386–402. (Cited on p. 226.)

Dennett, Daniel C.

1987 *The Intentional Stance*, MIT Press, Cambridge, Mass. (Cited on p. 276.)
1996 *Kinds of Minds: Toward an Understanding of Consciousness*, Basic Books, New York. (Cited on p. 268.)

DePaul, Michael R. and William M. Ramse y

1998 *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry*, Rowman & Littlefield, Lanham. (Cited on p. 328.)

Descartes, René

1633 "Treatise on Man," in Descartes [1985a], pp. 99–108. (Cited on p. 265.)
1641 "Meditations on First Philosophy," in Descartes [1985b], pp. 1–397. (Cited on pp. 41, 265.)
1649 "The Passions of the Soul," in Descartes [1985a], pp. 325–404. (Cited on p. 265.)
1964–76 *Oevres de Descartes*, ed. by Charles Adam and Paul Tannery, 11 volumes, Vrin/CNRS, Paris. (Cited on pp. 41, 265.)
1985a *The Philosophical Writings of Descartes: Volume 1*, ed. by John Cottingham, Robert Stoothoff, and Dugald Murdoch, Cambridge University Press, Cambridge. (Cited on p. 332.)
1985b *The Philosophical Writings of Descartes: Volume 2*, ed. by John Cottingham, Robert Stoothoff, and Dugald Murdoch, Cambridge University Press, Cambridge. (Cited on p. 332.)

Devitt, Michael

2008 "Resurrecting Biological Essentialism," *Philosophy of Science*, 75(3), pp. 344–82. (Cited on p. 303.)

Dieks, Dennis

2006 (ed.), *The Ontology of Spacetime*, Elsevier, Amsterdam. (Cited on p. 252.)
2008 (ed.), *The Ontology of Spacetime II*, Elsevier, Amsterdam. (Cited on p. 252.)

Dieks, Dennis, Wenceslao J. Gonzalez, Stephan Hartmann, Thomas Uebel, and Marcel Weber

2011 (eds.), *Explanation, Prediction, and Confirmation*, Springer, Dordrecht. (Cited on p. 338.)

Dobzhansky, Theodosius

1973 "Nothing in Biology Makes Sense Except in the Light of Evolution," *American Biology Teacher*, 35(3), pp. 125–9. (Cited on p. 296.)

Dorr, Cian

2005 "What we Disagree about when we Disagree about Ontology," in Kalderon [2005], ch. 8, pp. 234–86. (Cited on pp. 20, 174.)

Dorr, Cian and Gideon Rosen
2002 "Composition as Fiction," in Gale [2002], ch. 8, pp. 151–70. (Cited on p. 20.)
Dowe, Phil
2000 *Physical Causation*, Cambridge University Press, Cambridge. (Cited on p. 193.)
Dretske, Fred I.
1977 "Laws of Nature," *Philosophy of Science*, 44(2), pp. 248–68. (Cited on p. 189.)
Driesch, Hans
1914 *The History and Theory of Vitalism*, Macmillan and Co., London. (Cited on p. 268.)
Dummett, Michael
1960 "A Defense of Mctaggart's Proof of the Unreality of Time," *The Philosophical Review*, 69(4), pp. 497–504. (Cited on p. 214.)
1973 *Frege: Philosophy of Language*, Duckworth, London. (Cited on pp. 31, 100.)
1991 *The Logical Basis of Metaphysics*, Duckworth, London. (Cited on p. 102.)
2004 *Truth and the Past*, Columbia University Press, Chichester. (Cited on p. 226.)
Dunn, Jeffrey
2011 "Fried Eggs, Thermodynamics, and the Special Sciences," *The British Journal for the Philosophy of Science*, 62(1), pp. 71–98. (Cited on p. 172.)
Dupré, John
1993 *The Disorder of Things: Metaphysical Foundations of the Disunity of Science*, Harvard University Press, Cambridge, Mass. (Cited on pp. 266, 272.)
2010 "It Is Not Possible to Reduce Biological Explanations to Explanations in Chemistry and/or Physics," in Ayala *et al.* [2010], ch. 2, pp. 32–47. (Cited on pp. 272, 275, 299.)
2012 *Processes of Life: Essays in the Philosophy of Biology*, Oxford University Press, Oxford. (Cited on p. 272.)
2013 "Living Causes," *Aristotelian Society Supplementary Volume*, 87(1), pp. 19–37. (Cited on p. 272.)
Dyke, Heather
2002 "McTaggart and the Truth about Time," *Philosophy*, 50(Suppl.), pp. 137–52. (Cited on p. 214.)
Dyke, Heather and James Maclaurin
2013 "What Shall We Do with Analytic Metaphysics? A Response to McLeod and Parsons," *Australasian Journal of Philosophy*, 91(1), pp. 179–82. (Cited on p. 27.)
Earman, John
1986 *A Primer on Determinism*, D. Reidel. (Cited on p. 250.)
2008 "Reassessing the Prospects for a Growing Block Model of the Universe," *International Studies in the Philosophy of Science*, 22(2), pp. 135–64. (Cited on p. 243.)
Earman, John and John T. Roberts
1999 ""Ceteris Paribus", There Is No Problem of Provisos," *Synthese*, 118(3), pp. 439–78. (Cited on p. 139.)
2005a "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), pp. 1–22. (Cited on p. 172.)
2005b "Contact with the Nomic: A Challenge for Deniers of Humean Supervenience About Laws of Nature, Part II: The Epistemological Argument for Humean Supervenience," *Philosophy and Phenomenological Research*, 71(2), pp. 253–86. (Cited on p. 172.)
Eddon, Maya
2010 "Three Arguments From Temporary Intrinsics," *Philosophy and Phenomenological Research*, 81(3), pp. 605–19. (Cited on p. 207.)
Eklund, Matti
2008 "The Picture of Reality as an Amorphous Lump," in Sider *et al.* [2008], ch. 9.2, pp. 382–96. (Cited on pp. 31, 174.)

Eklund, Matti

2013 "Metaphysical Vagueness and Metaphysical Indeterminacy," *Metaphysica*, 14(2), pp. 165–79. (Cited on p. 307.)

Elder, Crawford L.

2005 *Real Natures and Familiar Objects*, The MIT Press. (Cited on p. 18.)

2011 *Familiar Objects and Their Shadows*, Cambridge University Press, Cambridge. (Cited on pp. 18, 30, 31, 42, 62, 115, 284.)

2013a "Millikan, Realism, and Sameness," in Ryder *et al.* [2013], ch. 8, pp. 155–71. (Cited on p. 286.)

2013b "On the Reality and Causal Efficacy of Familiar Objects," *Philosophia*, 41(3), pp. 737–49. (Cited on p. 284.)

Elga, Adam

2001 "Statistical Mechanics and the Asymmetry of Counterfactual Dependence," *Philosophy of Science*, 68(s1), pp. 313–24. (Cited on p. 186.)

Ellis, Brian D.

2001 *Scientific Essentialism*, Cambridge University Press. (Cited on p. 198.)

2010 "Causal Powers and Categorical Properties," in Marmodoro [2010], pp. 133–42. (Cited on p. 198.)

Ellis, Brian D. and Caroline Lierse

1994 "Dispositional Essentialism," *Australasian Journal of Philosophy*, 72(1), pp. 27–45. (Cited on p. 198.)

Ellis, Fiona

2005 *Concepts and Reality in the History of Philosophy: Tracing a Philosophical Error from Locke to Bradley*, Routledge, London. (Cited on pp. 8, 17, 266, 269, 317.)

Ellis, George F. R.

2006 "Physics in the Real Universe: Time and Spacetime," *General Relativity and Gravitation*, 38, pp. 1797–824. (Cited on pp. 248, 253.)

Engel, Pascal

2001 "The False Modesty of the Identity Theory of Truth," *International Journal of Philosophical Studies*, 9(4), pp. 441–58. (Cited on p. 317.)

Etchemendy, John

1983 "The Doctrine of Logic as Form," *Linguistics and Philosophy*, 6(3), pp. 319–34. (Cited on pp. 90, 91.)

1988 "Tarski on Truth and Logical Consequence," *The Journal of Symbolic Logic*, 53(1), pp. 51–79. (Cited on pp. 90, 91.)

1990 *The Concept of Logical Consequence*, Harvard University Press, Cambridge, Mass. (Cited on pp. 90, 91, 94, 96, 98, 99.)

Evans, Gareth

1976 "Semantic Structure and Logical Form," in McDowell *et al.* [1976], pp. 199–222. (Cited on p. 100.)

1982 *The Varieties of Reference*, ed. by John Mcdowell, Clarendon Press, Oxford. (Cited on pp. 45, 112.)

1985a *Collected Papers*, Clarendon Press, Oxford. (Cited on p. 334.)

1985b "Does Tense Logic Rest Upon a Mistake?" In Evans [1985a], ch. 12, pp. 343–63. (Cited on p. 242.)

Everett, Hugh

1957 ""Relative State" Formulation of Quantum Mechanics," *Review of Modern Physics*, 29, pp. 454–62. (Cited on p. 253.)

Falvey, Kevin

2010 "The View From Nowhen: The Mctaggart-Dummett Argument for the Unreality of Time," *Philosophia*, 38(2). (Cited on p. 214.)

Fara, Michael
2005 "Dispositions and Habituals," *Noûs*, 39(1), pp. 43–82. (Cited on p. 197.)
Feferman, Solomon
1999 "Logic, Logics, and Logicism," *Notre Dame Journal of Formal Logic*, 40, pp. 31–54. (Cited on p. 101.)
Feigl, Herbert and May Brodbeck
1953 *Readings in the Philosophy of Science*, Appleton Century Crofts, New York. (Cited on p. 349.)
Fine, Kit
1975 "Critical Notice: *Counterfactuals* by D. Lewis," *Mind*, 84(335), pp. 451–8. (Cited on p. 186.)
1989 "The Problem of De Re Modality," in Fine [2005a], pp. 40–104. (Cited on pp. 79, 107.)
1990 "Quine on Quantifying In," in Fine [2005a], pp. 105–32. (Cited on p. 79.)
1994a "Compounds and Aggregates," *Noûs*, 28(2), pp. 137–58. (Cited on p. 24.)
1994b "Essence and Modality: The Second Philosophical Perspectives Lecture," *Philosophical Perspectives*, 8, pp. 1–16. (Cited on pp. 23, 44, 62, 65, 130.)
1999 "Things and Their Parts," *Midwest Studies in Philosophy*, 23(1), pp. 61–74. (Cited on p. 24.)
2001 "The Question of Realism," *Philosopher's Imprint*, 1(2), pp. 1–30. (Cited on pp. 2, 44, 49.)
2002 "The Problem of Possibilia," in Fine [2005a], pp. 214–34. (Cited on p. 95.)
2003 "The Non-Identity of a Material Thing and Its Matter," *Mind*, 112(446), pp. 195–234. (Cited on pp. 23, 44.)
2005a *Modality and Tense: Philosophical Papers*, Oxford University Press. (Cited on p. 335.)
2005b "Necessity and Non-Existence," in Fine [2005a], ch. 9, pp. 321–354. (Cited on pp. 108, 109.)
2005c "Reference, Essence and Identity," in Fine [2005a], pp. 19–39. (Cited on pp. 44, 79.)
2005d "Tense and Reality," in Fine [2005a], pp. 261–320. (Cited on pp. 214, 217–220, 222, 225, 247, 253.)
2006 "In Defense of Three-Dimensionalism," *Journal of Philosophy*, 103(12), pp. 699–714. (Cited on p. 208.)
Fiocco, Marcello Oreste
2007 "A Defense of Transient Presentism," *American Philosophical Quarterly*, 44(3), pp. 191–212. (Cited on p. 226.)
Forrest, Peter
2006 "General Facts, Physical Necessity and the Metaphysics of Time," in Zimmerman [2006], ch. 6, pp. 137–53. (Cited on p. 243.)
Frege, Gottlob
1879 *Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens*, Louis Nebert, Halle a. S. (Cited on pp. 111, 114.)
1892a "Über Begriff und Gegenstand," *Vierteljahresschrift für wissenschaftliche Philosophie*, 16, pp. 192–205. (Cited on pp. 111, 114.)
1892b "Über Sinn und Bedeutung," *Zeitschrift für Philosophie und philosophische Kritik*, 100, pp. 25–50. (Cited on pp. 111, 114.)
1978/1884 *The Foundations of Arithmetic: A Logico-Mathematical Enquiry Into the Concept of Number*, Translated by John L. Austin, Northwestern University Press, Evanston, Illinois. (Cited on p. 3.)
French, Steven
2011 *Identity and Individuality in Quantum Theory*, ed. by Edward N. Zalta, http://plato.stanford.edu/archives/sum2011/entries/qt-idind/. (Cited on p. 273.)
French, Steven and Décio Krause
2006 *Identity in Physics: A Historical, Philosophical, and Formal Analysis*, Oxford University Press, New York. (Cited on p. 273.)
French, Steven and Michael Redhead
1988 "Quantum Physics and the Identity of Indiscernibles," *British Journal for the Philosophy of Science*, 39(2), pp. 233–46. (Cited on p. 273.)

French, Steven and Juha Saatsi
2011 (eds.), *The Continuum Companion to Philosophy of Science*, Continuum, New York. (Cited on p. 330.)
Friebe, Cord
2012a "Twins' Paradox and Closed Timelike Curves: The Role of Proper Time and the Presentist View on Spacetime," *Journal for General Philosophy of Science*, 43(2), pp. 313–26. (Cited on p. 252.)
2012b *Zeit – Wirklichkeit – Persistenz: Eine präsentistische Deutung der Raumzeit*, mentis, Paderborn. (Cited on p. 252.)
Frisch, Mathias
2011 "From Arbuthnot to Boltzmann: The Past Hypothesis, the Best System, and the Special Sciences," *Philosophy of Science*, 78(5), pp. 1001–11. (Cited on p. 173.)
Gale, Richard M.
2002 (ed.), *The Blackwell Guide to Metaphysics*, Wiley-Blackwell. (Cited on p. 333.)
Geach, Peter T.
1967 "Some Problems About Time," in *Proceedings of the British Academy*, Oxford University Press, London, vol. LI. (Cited on p. 216.)
1972 *Logic Matters*, University of California Press, Berkeley. (Cited on p. 163.)
1980 *Reference and Generality*, 3rd edition, Cornell University Press, Ithaca, New York. (Cited on pp. 24, 66.)
Gentzen, Gerhard
1969 *The Collected Papers of Gerhard Gentzen*, North-Holland, Amsterdam. (Cited on p. 102.)
Ghiselin, Michael T.
1974 "A Radical Solution to the Species Problem," *Systematic Zoology*, 23, pp. 536–44. (Cited on p. 303.)
1997 *Metaphysics and the Origin of Species*, State University of New York Press, Albany. (Cited on p. 303.)
Gibbard, Alan
1975 "Contingent Identity," *Journal of Philosophical Logic*, 4(2), pp. 187–221. (Cited on p. 24.)
Giberman, Daniel
2014 "Tropes in Space," *Philosophical Studies*, 167(2), pp. 453–472. (Cited on p. 70.)
Goldfarb, Warren
1985 "Kripke on Wittgenstein on Rules," *The Journal of Philosophy*, 82(9), pp. 471–88. (Cited on p. 283.)
Goldschmidt, Richard B.
1940 *The Material Basis of Evolution*, Yale University Press, New Haven. (Cited on p. 303.)
Goodman, Nelson
1955 *Fact, Fiction and Forecast*, Harvard University Press, Cambridge, Mass. (Cited on pp. 178, 185, 197.)
1978 *Ways of Worldmaking*, Hackett Publishing Company, Indianapolis. (Cited on p. 29.)
Gould, Stephen J.
2002 *The Structure of Evolutionary Theory*, Belknap Press, Cambridge, Mass. (Cited on pp. 298, 303.)
Greco, John and Ruth Groff
2013 (eds.), *Powers and Capacities in Philosophy: The New Aristotelianism*, Routledge, New York. (Cited on p. 23.)
Griesemer, James R.
2000 "Reproduction and the Reduction of Genetics," in Beurton *et al.* [2000], ch. 11, pp. 240–85. (Cited on p. 295.)

Gunther, York H.
2003 (ed.), *Essays on Nonconceptual Content*, The MIT Press, Cambridge, Mass. (Cited on p. 112.)
Gutting, Gary
2009 *What Philosophers Know: Case Studies in Recent Analytic Philosophy*, Cambridge University Press, Cambridge. (Cited on p. 116.)
Hacking, Ian
1979 "What is Logic?" *Journal of Philosophy*, 76(6), pp. 285–319. (Cited on p. 102.)
Haddock, Adrian and Fiona Macpherson
2008 (eds.), *Disjunctivism: Perception, Action, Knowledge*, Oxford University Press, New York. (Cited on p. 41.)
Haldane, John S.
1931 *The Philosophical Basis of Life*, Doubleday, Doran, Garden City. (Cited on p. 269.)
Halpern, Joseph Y. and Judea Pearl
2005 "Causes and Explanations: A Structural-Model Approach. Part I: Causes," *The British Journal for the Philosophy of Science*, 56(4), pp. 843–87. (Cited on p. 188.)
Haslanger, Sally
1989 "Endurance and Temporary Intrinsics," *Analysis*, 49, pp. 119–25. (Cited on p. 208.)
Hawley, Katherine
2002 *How Things Persist*, Oxford University Press, New York. (Cited on pp. 127, 205, 207, 208, 245, 246.)
Hegel, Georg Wilhelm Friedrich
1969/1816 *Science of Logic*, Translated from the German by A.V. Miller, Humanity Books, Amherst. (Cited on p. 18.)
1977/1807 *Phenomenology of Spirit*, Translated from the German by A.V. Miller, Oxford University Press, New York. (Cited on p. 18.)
Heil, John
2003 *From an Ontological Point of View*, Clarendon Press, Oxford. (Cited on pp. 23, 127, 198.)
2005 "Dispositions," *Synthese*, 144(3), pp. 343–56. (Cited on pp. 127, 198.)
Heller, Mark
1990 *The Ontology of Physical Objects: Four-Dimensional Hunks of Matter*, Cambridge University Press, Cambridge. (Cited on pp. 26, 78.)
1996 "Against Metaphysical Vagueness," *Noûs*, 30, pp. 177–85. (Cited on p. 307.)
Hempel, Carl G.
1966 *Philosophy of Natural Science*, Prentice-Hall, Englewood Cliffs, N.J. (Cited on p. 266.)
Hinchliff, Mark
2000 "A Defense of Presentism in a Relativistic Setting," *Philosophy of Science*, 67(3), pp. 575–86. (Cited on pp. 216, 252.)
Hirsch, Eli
2002a "Against Revisionary Ontology," *Philosophical Topics*, 30, pp. 103–27. (Cited on pp. 18, 29.)
2002b "Quantifier Variance and Realism," *Philosophical Issues*, 12(1), pp. 51–73. (Cited on pp. 18, 29.)
2005 "Physical-Object Ontology, Verbal Disputes, and Common Sense," *Philosophy and Phenomenological Research*, 70(1), pp. 67–97. (Cited on pp. 18, 29.)
2008 "Language, Ontology, and Structure," *Noûs*, 42(3), pp. 509–28. (Cited on pp. 18, 29.)
2009 "Ontology and Alternative Languages," in Chalmers *et al.* [2009], ch. 8, pp. 231–59. (Cited on p. 208.)
Hitchcock, Christopher and James Woodward
2003 "Explanatory Generalizations, Part II: Plumbing Explanatory Depth," *Noûs*, 37(2), pp. 181–99. (Cited on p. 187.)

Hodes, Harold
2004 "On the Sense and Reference of a Logical Constant," *Philosophical Quarterly*, 54(214), pp. 134–65. (Cited on p. 102.)

Hoerl, Christoph, Teresa McCormack, and Sarah Beck
2012 (eds.), *Understanding Counterfactuals, Understanding Causation: Issues in Philosophy and Psychology*, Oxford University Press, New York. (Cited on p. 186.)

Horgan, Terence and Matjaž Potrč
2008 *Austere Realism: Contextual Semantics Meets Minimal Ontology*, The MIT Press, Cambridge, Mass. (Cited on pp. 20, 25, 361.)

Hornsby, Jennifer
1997 "The Presidential Address: Truth: The Identity Theory," *Proceedings of the Aristotelian Society*, 97, pp. 1–24. (Cited on p. 317.)

Hull, David L.
1965a "The Effect of Essentialism on Taxonomy—-Two Thousand Years of Stasis (I)," *British Journal for the Philosophy of Science*, 15(60), pp. 314–26. (Cited on p. 301.)
1965b "The Effect of Essentialism on Taxonomy—Two Thousand Years of Stasis (II)," *British Journal for the Philosophy of Science*, 16(61), pp. 1–18. (Cited on p. 301.)
1974 *The Philosophy of Biological Science*, Prentice-Hall, Englewood Cliffs, N.J. (Cited on p. 266.)

Hume, David
2007/1748 *An Enquiry Concerning Human Understanding*, Ed. by Peter Millican, Oxford University Press, New York. (Cited on pp. 169, 170, 185.)
2010/1740 *A Treatise of Human Nature*, Oxford University Press, Oxford. (Cited on p. 170.)

Husserl, Edmund
1970/1954 *The Crisis of European Sciences and Transcendental Phenomenology*, Translated from the German by D. Carr, Northwestern University Press, Evanston, Illinois. (Cited on p. 18.)

Hüttemann, Andreas
2011 "Comparing Part-Whole Reductive Explanations in Biology and Physics," in Dieks *et al.* [2011], pp. 183–202. (Cited on p. 275.)
2013 "A Disposition-Based Process Theory of Causation," in Mumford *et al.* [2013], ch. 5, pp. 101–22. (Cited on p. 198.)

Huxley, Julian S.
1942 *Evolution: The Modern Synthesis*, Allen and Unwin, London. (Cited on p. 303.)

Jablonka, Eva and Marion J. Lamb
2005 *Evolution in Four Dimensions: Genetic, Epigenetic, Behavioral, and Symbolic Variation in the History of Life*, MIT Press, Cambridge, Mass. (Cited on p. 307.)
2007 "Bridging the Gap: The Developmental Aspects of Evolution," *Behavioral and Brain Sciences*, 30(4), pp. 378–89. (Cited on p. 307.)
2008 "Soft Inheritance: Challenging the Modern Synthesis," *Genetics and Molecular Biology*, 31(2), pp. 389–95. (Cited on p. 307.)

Jablonka, Eva and Gal Raz
2009 "Transgenerational Epigenetic Inheritance: Prevalence, Mechanisms, and Implications for the Study of Heredity and Evolution," *Quarterly Review of Biology*, 84, pp. 131–76. (Cited on p. 307.)

Jackson, Frank
1977 "A Causal Theory of Counterfactuals," *Australasian Journal of Philosophy*, 55(1), pp. 3–21. (Cited on p. 186.)

James, William
1890 *The Principles of Psychology*, Harvard University Press, Cambridge, Mass. (Cited on p. 1.)

Jenkins, Carrie S.
2010 "What Is Ontological Realism?" *Philosophy Compass*, 5(10), pp. 880–90. (Cited on p. 22.)

Johnston, Mark
1987 "Is There a Problem About Persistence?" *Proceedings of the Aristotelian Society Supplementary Volume*, 61, pp. 107–35. (Cited on pp. 164, 208.)
1992 "How to Speak of the Colors," *Philosophical Studies*, 68(3), pp. 221–63. (Cited on p. 197.)
Jubien, Michael
1993 *Ontology, Modality and the Fallacy of Reference*, Cambridge University Press, Cambridge. (Cited on p. 31.)
1996 "The Myth of Identity Conditions," *Noûs*, 30, pp. 343–56. (Cited on pp. 68, 73, 74.)
Kalderon, Mark Eli
2005 (ed.), *Fictionalism in Metaphysics*, Oxford University Press, New York. (Cited on p. 332.)
Kant, Immanuel
1790 *Kritik der Urteilskraft*, Lagarde & Friedrich, Berlin & Libau. (Cited on p. 276.)
Keefe, Rosanna and Peter Smith
1997 (eds.), *Vagueness: A Reader*, MIT Press, Cambridge, Mass. (Cited on p. 307.)
Keinänen, Markku
2011 "Tropes – The Basic Constituents of Powerful Particulars?" *dialectica*, 65(3), pp. 419–50. (Cited on p. 70.)
Keinänen, Markku and Jani Hakkarainen
2013 "The Problem of Trope Individuation: A Reply to Lowe," *Erkenntnis*, pp. 1–15. (Cited on p. 70.)
Keller, Evelyn Fox
2010 "It Is Possible to Reduce Biological Explanations to Explanations in Chemistry and/or Physics," in Ayala *et al.* [2010], ch. 1, pp. 19–31. (Cited on p. 299.)
Kemp Smith, Norman
1918 *Commentary to Kant's Critique of Pure Reason*, The Macmillan Press Ltd., London. (Cited on p. 30.)
Kitcher, Philip S.
1984 "1953 and All That. A Tale of Two Sciences," *The Philosophical Review*, 93(3), pp. 335–71. (Cited on p. 266.)
Knobe, Joshua M. and Shaun Nichols
2008 *Experimental Philosophy*, Oxford University Press, New York. (Cited on p. 118.)
2013 (eds.), *Experimental Philosophy, Volume 2*, Oxford University Press, New York. (Cited on p. 118.)
Koslicki, Kathrin
2004 "Constitution and Similarity," *Philosophical Studies*, 117(3), pp. 327–63. (Cited on p. 24.)
2008 *The Structure of Objects*, Oxford University Press, New York. (Cited on p. 24.)
Kreines, James
2008 "The Logic of Life: Hegel's Philosophical Defense of Teleological Explanation of Living Beings," in Beiser [2008], ch. 13, pp. 344–77. (Cited on p. 276.)
Kreisel, Georg
1967 "Informal Rigour and Completeness Proofs," in Lakatos [1967], pp. 138–57. (Cited on pp. 91, 99, 104.)
Kripke, Saul A.
1972 *Naming and Necessity*, Harvard University Press, Cambridge, Mass. (Cited on pp. 60, 118, 166, 280.)
1982 *Wittgenstein on Rules and Private Language*, Blackwell Publishing, Oxford. (Cited on p. 283.)
Ladyman, James
2002 "Science, Metaphysics, and Structural Realism," *Philosophica*, 67, pp. 57–76. (Cited on p. 23.)

Ladyman, James
2012 "Science, Metaphysics and Method," *Philosophical Studies*, 160(1), pp. 31–51. (Cited on pp. 3, 18, 27.)

Ladyman, James, Don Ross, David Spurrett, and John Collier
2007 *Every Thing Must Go: Metaphysics Naturalized*, Oxford University Press, New York. (Cited on pp. 3, 18, 23, 24, 27, 72.)

Lakatos, Imre
1967 (ed.), *Problems in the Philosophy of Mathematics*, North-Holland, Amsterdam. (Cited on p. 339.)

Landry, Elaine and Dean Rickles
2012 (eds.), *Structural Realism: Structure, Object, and Causality*, Springer. (Cited on p. 23.)

Lange, Marc
1993 "Natural Laws and the Problem of Provisos," *Erkenntnis*, 38(2), pp. 233–248. (Cited on p. 139.)

Laplace, Pierre-Simon
1995/1814 *A Philosophical Essay on Probabilities*, Translated from the fifth French edition of 1825 by Andrew I. Dale, Springer Verlag, New York. (Cited on p. 147.)

Laycock, Henry
2006 *Words without Objects: Semantics, Ontology, and Logic for Non-Singularity*, Oxford University Press, New York. (Cited on pp. 114, 138.)

Levit, Georgy S. and Kay Meister
2006 "The History of Essentialism vs. Ernst Mayr's "Essentialism Story": A Case Study of German Idealistic Morphology," *Theory in Biosciences*, 124(3–4), pp. 281–307. (Cited on p. 302.)

Lewis, David K.
1968 "Counterpart Theory and Quantified Modal Logic," *Journal of Philosophy*, 65(5), pp. 113–26. (Cited on pp. 114, 166.)
1973a "Causation," *The Journal of Philosophy*, 70(17), pp. 556–67. (Cited on p. 185.)
1973b *Counterfactuals*, Blackwell Publishing, Oxford. (Cited on pp. 114, 166, 172, 185.)
1979a "Attitudes *De Dicto* and *De Se*," *The Philosophical Review*, 88(4), pp. 513–43. (Cited on p. 215.)
1979b "Counterfactual Dependence and Time's Arrow," *Noûs*, 13(4), pp. 455–76. (Cited on p. 186.)
1983 "New Work for a Theory of Universals," *Australasian Journal of Philosophy*, 61(4), pp. 343–77. (Cited on pp. 173, 190.)
1986a *On the Plurality of Worlds*, Blackwell Publishing. (Cited on pp. 20, 22, 77, 81, 95, 114, 127, 160, 166, 173, 206–208, 228.)
1986b *Philosophical Papers: Volume II*, Oxford University Press, New York. (Cited on pp. 159, 165, 173, 186, 320, 373.)
1991 *Parts of Classes*, Blackwell Publishing, Oxford. (Cited on pp. 22, 77, 127, 162.)
1994 "Humean Supervenience Debugged," *Mind*, 103(412), pp. 473–90. (Cited on pp. 165, 172, 173, 187.)
1997 "Finkish Dispositions," *The Philosophical Quarterly*, 47(187), pp. 143–58. (Cited on p. 197.)
2000 "Causation as Influence," *The Journal of Philosophy*, 97(4), pp. 182–97. (Cited on p. 186.)
2002 "Tensing the Copula," *Mind*, 111(441), pp. 1–13. (Cited on pp. 160, 247.)
2004a "Causation as Influence," in Collins *et al.* [2004], ch. 3, pp. 75–106. (Cited on p. 186.)
2004b "Void and Object," in Collins *et al.* [2004], ch. 10, pp. 277–90. (Cited on p. 186.)

Locke, John
1790 *An Essay Concerning Human Understanding*, The Baffet, London. (Cited on p. 27.)

Loewer, Barry
1996 "Humean Supervenience and Laws of Nature," *Philosophical Topics*, 24, pp. 101–27. (Cited on p. 173.)

Loewer, Barry

2007a "Counterfactuals and the Second Law," in Price *et al.* [2007], ch. 11, pp. 293–326. (Cited on p. 173.)

2007b "Laws and Natural Properties," *Philosophical Topics*, 35(1/2), pp. 313–28. (Cited on p. 172.)

2009 "Why Is There Anything Except Physics?" *Synthese*, 170(2), pp. 217–33. (Cited on p. 172.)

2012 "Two Accounts of Laws and Time," *Philosophical Studies*, 160(1), pp. 115–37. (Cited on p. 172.)

Loux, Michael J. and Dean W. Zimmerman

2003 *The Oxford Handbook of Metaphysics*, Oxford University Press, New York. (Cited on pp. 331, 341.)

Lowe, E. J.

1987 "Lewis on Perdurance versus Endurance," *Analysis*, 47, pp. 152–4. (Cited on p. 208.)

1988 "The Problems of Intrinsic Change: Rejoinder to Lewis," *Analysis*, 48(2), pp. 72–7. (Cited on p. 208.)

1991 "One-Level versus Two-Level Identity Criteria," *Analysis*, 51(4), pp. 192–94. (Cited on pp. 66, 67.)

1992 "Mctaggart's Paradox Revisited," *Mind*, 101(402), pp. 323–326. (Cited on p. 214.)

1998 *The Possibility of Metaphysics: Substance, Identity, and Time*, Oxford University Press. (Cited on p. 216.)

2003 "Indviduation," in Loux *et al.* [2003], ch. 3, pp. 75–95. (Cited on p. 59.)

2006 *The Four-Category Ontology: A Metaphysical Foundation for Natural Science*, Oxford University Press, New York. (Cited on pp. 23, 60, 70, 141, 198.)

2008a "New Directions in Metaphysics and Ontology," *Axiomathes*, 18(3), pp. 273–88. (Cited on p. 44.)

2008b *Personal Agency: The Metaphysics of Mind and Action*, Oxford University Press. (Cited on p. 23.)

2009 *More Kinds of Being: A Further Study of Individuation, Identity, and the Logic of Sortal Terms*, Wiley-Blackwell, Chichester. (Cited on pp. 23, 44, 60, 62, 66, 132, 136, 141.)

2013a "Ontological Vagueness, Existence Monism and Metaphysical Realism," *Metaphysica*, 14(2), pp. 265–74. (Cited on p. 307.)

2013b "Presentism and Relativity: No Conflict," in Ciuni *et al.* [2013], pp. 133–52. (Cited on p. 253.)

Lowe, E. J. and A. Rami

2008 (eds.), *Truth and Truth-Making*, McGill-Queen's University Press, Montreal and Kingston. (Cited on p. 38.)

MacBride, Fraser

2004 "Review of *Resemblance Nominalism* by Rodriguez-Pereyra," *Notre Dame Philosophical Reviews*. (Cited on p. 28.)

MacDonald, Graham

1979 (ed.), *Perception and Identity: Essays Presented to A.J. Ayer, With His Replies*, Cambridge University Press, Cambridge. (Cited on p. 328.)

MacFarlane, John

2000 *What Does it Mean to Say that Logic is Formal?* Ph. D. Dissertation, Pittsburgh. (Cited on pp. 96, 318.)

2009 *Logical Constants*, ed. by Edward N. Zalta, http://plato.stanford.edu/archives/fall2009/entries/logical-constants/. (Cited on pp. 100, 101, 112, 116, 318.)

Mackie, Penelope

2006 *How Things Might Have Been: Individuals, Kinds, and Essential Properties*, Oxford University Press, New York. (Cited on p. 79.)

Maclaurin, James and Heather Dyke

2012 "What is Analytic Metaphysics For?" *Australasian Journal of Philosophy*, 90(2), pp. 291–306. (Cited on p. 27.)

Malpass, Alex and Jacek Wawer
2012 "A Future for the Thin Red Line," *Synthese*, 188(1), pp. 117–42. (Cited on p. 137.)

Markosian, Ned
1993 "How Fast Does Time Pass?" *Philosophy and Phenomenological Research*, 53(4), pp. 829–44. (Cited on p. 221.)
1995 "The Open Past," *Philosophical Studies*, 79(1), pp. 95–105. (Cited on p. 226.)
2004 "A Defense of Presentism," in Zimmerman [2004], ch. 3, pp. 47–82. (Cited on p. 216.)
2013 "The Truth About the Past and the Future," in Correia *et al.* [2013], pp. 127–42. (Cited on pp. 216, 226.)

Marmodoro, Anna
2009 "'Do powers need powers to make them powerful?' From Pandispositionalism to Aristotle," *History of Philosophy Quarterly*, 26(4). (Cited on pp. 72, 82, 83, 198, 300.)
2010 *The Metaphysics of Powers: Their Grounding and their Manifestations*, Routledge. (Cited on pp. 23, 334.)

Martin, C. B.
1994 "Dispositions and Conditionals," *The Philosophical Quarterly*, 44(174), pp. 1–8. (Cited on p. 197.)
1997 "On the Need for Properties: The Road to Pythagoreanism and Back," *Synthese*, 112(2), pp. 193–231. (Cited on pp. 127, 198.)

Martin, C. B. and John Heil
1999 "The Ontological Turn," *Midwest Studies in Philosophy*, 23(1), pp. 34–60. (Cited on pp. 127, 198.)

Martin-Löf, Per
1987 "Truth of a Proposition, Evidence of a Judgement, Validity of a Proof," *Synthese*, 73(3), pp. 407–20. (Cited on p. 103.)

Mayden, Richard L.
1997 "A Hierarchy of Species Concepts: The Denouement in the Saga of the Species Problem," in Claridge *et al.* [1997], ch. 19, pp. 381–424. (Cited on p. 305.)

Mayr, Ernst
1940 "Speciation Phenomena in Birds," *American Naturalist*, 74, pp. 249–78. (Cited on p. 304.)
1942 *Systematics and the Origin of Species From the Viewpoint of a Zoologist*, Columbia University Press, New York. (Cited on p. 305.)
1970 *Populations, Species, and Evolution: An Abridgment of Animal Species and Evolution*, Belknap Press, Cambridge, Mass. (Cited on p. 304.)

McCall, Storrs
1976 "Objective Time Flow," *Philosophy of Science*, 43(3), pp. 337–62. (Cited on p. 249.)
1994 *A Model of the Universe*, Clarendon Press, Oxford. (Cited on p. 216.)
2012 "The Origin of Life and the Definition of Life," in Tahko [2012], ch. 11, pp. 174–86. (Cited on p. 261.)

McCarthy, Timothy
1981 "The Idea of a Logical Constant," *Journal of Philosophy*, 78, pp. 499–523. (Cited on pp. 100, 101.)

McDowell, John H.
1989 "Wittgenstein and the Inner World," *The Journal of Philosophy*, 86(11), pp. 643–4. (Cited on p. 52.)
1990 "Peacocke and Evans on Demonstrative Content," *Mind*, 99(394), pp. 255–66. (Cited on p. 112.)
1994a *Mind and World*, Harvard University Press, Cambridge, Mass. (Cited on pp. 17, 41, 52, 112, 314, 317.)
1994b "Nonconceptual Content," in pp. 75–90. (Cited on p. 112.)
1994c "The Content of Perceptual Experience," *Philosopical Quarterly*, 44(175), pp. 190–205. (Cited on p. 112.)

McDowell, John H.

2005 "The True Modesty of an Identity Conception of Truth: A Note in Response to Pascal Engel," *International Journal of Philosophical Studies*, 13(1), pp. 83–8. (Cited on p. 317.)

2009a *Having the World in View: Essays on Kant, Hegel, and Sellars*, Harvard University Press, Cambridge, Mass. (Cited on p. 17.)

2009b *The Engaged Intellect: Philosophical Essays*, Harvard University Press, Cambridge, Mass. (Cited on pp. 17, 52, 343.)

2009c "Towards Rehabilitating Objectivity," in McDowell [2009b], ch. 12, pp. 204–24. (Cited on p. 52.)

2011 *Perception as a Capacity for Knowledge*, Marquette University Press, Milwaukee, WI. (Cited on pp. 41, 112.)

McDowell, John H. and Gareth Evans

1976 (eds.), *Truth and Meaning: Essays in Semantics*, Oxford University Press, New York. (Cited on p. 334.)

McDowell, John H. and Marcus Willaschek

2000 *Reason and Nature: Lecture and Colloquium in Münster 1999*, LIT, Münster. (Cited on p. 52.)

McGee, Vann

1996 "Logical Operations," *Journal of Philosophical Logic*, 25, pp. 567–80. (Cited on p. 101.)

McOuat, Gordon R.

2001 "From Cutting Nature at its Joints to Measuring It: New Kinds and New Kinds of People in Biology," *Studies in History and Philosophy of Science Part A*, 32(4), pp. 613–45. (Cited on p. 302.)

McTaggart, John M. E.

1908 "The Unreality of Time," *Mind*, 17(68), pp. 457–74. (Cited on pp. 212, 213.)

1927 *The Nature of Existence – Volume II*, Cambridge University Press, Cambridge. (Cited on p. 163.)

Mellor, David H.

1981 *Real Time*, Cambridge University Press, Cambridge. (Cited on pp. 163, 164, 207, 214.)

1998 *Real Time II*, Routledge, London. (Cited on pp. 164, 207, 208, 214, 215.)

Menzies, Peter and Huw Price

1993 "Causation as a Secondary Quality," *The British Journal for the Philosophy of Science*, 44(2), pp. 187–203. (Cited on p. 187.)

Merricks, Trenton

1994 "Endurance and Indiscernibility," *The Journal of Philosophy*, 91(4), pp. 165–84. (Cited on pp. 208, 211, 212.)

1999 "Persistence, Parts, and Presentism," *Noûs*, 33(3), pp. 421–38. (Cited on p. 216.)

2001 *Objects and Persons*, Oxford University Press. (Cited on p. 24.)

2003 "The End of Counterpart Theory," *The Journal of Philosophy*, 100(10), pp. 521–49. (Cited on p. 95.)

2006 "Goodbye Growing Block," in Zimmerman [2006], ch. 4, pp. 103–10. (Cited on p. 243.)

Mertz, Donald W.

2001 "Individuation and Instance Ontology," *Australasian Journal of Philosophy*, 79(1), pp. 45–61. (Cited on p. 79.)

2003 "Against Bare Particulars. A Response to Moreland and Pickavance," *Australasian Journal of Philosophy*, 81(1), pp. 14–20. (Cited on p. 79.)

Mill, John S.

1865 *An Examination of Sir William Hamilton's Philosophy*, Longman, London. (Cited on p. 170.)

1884 *A System of Logic*, Longman, London. (Cited on p. 302.)

Millgram, Elijah

2009 "Lewis's Epicycles, Possible Worlds, and the Mysteries of Modality," Unpublished manuscript. (Cited on p. 136.)

Millikan, Ruth Garrett
n.d. *Unicepts, Language, and Natural Information*, Oxford University Press, New York. (Cited on pp. 283–286.)
1984 *Language, Thought, and Other Biological Categories*, The MIT Press, Cambridge, Mass. (Cited on pp. 282, 287–290, 293–295, 309, 379.)
1989 "In Defense of Proper Functions," *Philosophy of Science*, 56(2), pp. 288–302. (Cited on pp. 290, 291, 294.)
1990 "Truth Rules, Hoverflies, and the Kripke-Wittgenstein Paradox," *The Philosophical Review*, 99(3), pp. 323–53. (Cited on p. 283.)
1993 *White Queen Psychology and Other Essays for Alice*, The MIT Press, Cambridge, Mass. (Cited on p. 282.)
1996 "On Swampkinds," *Mind and Language*, 11(1), pp. 103–17. (Cited on p. 291.)
1999 "Historical Kinds and the "Special Sciences"," *Philosophical Studies*, 95(1–2), pp. 45–65. (Cited on p. 285.)
2000 *On Clear and Confused Ideas*, Cambridge University Press, New York. (Cited on pp. 280–283.)
2004 *Varieties of Meaning*, The MIT Press, Cambridge, Mass. (Cited on pp. 281, 282.)
2005 *Language: A Biological Model*, Clarendon Press, Oxford. (Cited on p. 282.)
2010 "On Knowing the Meaning; With a Coda on Swampman," *Mind*, 119(473), pp. 43–81. (Cited on pp. 281, 291.)
2012 "Are There Mental Indexicals and Demonstratives?" *Philosophical Perspectives*, 26(1), pp. 217–34. (Cited on p. 279.)
2013 "Afterword," in Ryder *et al.* [2013], p. 281. (Cited on p. 279.)
2014 "An Epistemology for Phenomenology?" In R. Brown [2014], ch. 2, pp. 13–26. (Cited on pp. 279, 280.)

Mocek, Reinhard
1974 *Wilhelm Roux, Hans Driesch: zur Geschichte der Entwicklungsphysiologie der Tiere ("Entwicklungsmechanik")*, Fischer, Jena. (Cited on p. 259.)

Molnar, George
2003 *Powers: A Study in Metaphysics*, Oxford University Press, New York. (Cited on pp. 23, 127, 198.)

Moreland, James P.
1998 "Theories of Individuation: A Reconsideration of Bare Particularism," *Pacific Philosophical Quarterly*, 79(3), pp. 251–63. (Cited on p. 79.)
2001 *Universals*, McGill-Queen's University Press, Montreal and Kingston. (Cited on p. 79.)

Mulder, Jesse M.
2010 *Summing Things Up*, MA thesis, Utrecht University, http://igitur-archive.library.uu.nl/student-theses/2010-0302-200222/UUindex.html. (Cited on pp. 77, 78, 162.)
2012 "What Generates the Realism/Anti-Realism Dichotomy?" *Philosophica*, 84, pp. 49–80. (Cited on p. 17.)
2013 "The Essentialist Inference," *Australasian Journal of Philosophy*, 91(4), pp. 755–69. (Cited on pp. 59, 68.)

Muller, F. A. and Simon Saunders
2008 "Discerning Fermions," *British Journal for the Philosophy of Science*, 59(3), pp. 499–548. (Cited on p. 273.)

Muller, F. A. and M. P. Seevinck
2009 "Discerning Elementary Particles," *Philosophy of Science*, 76(2), pp. 179–200. (Cited on p. 273.)

Müller, Thomas
2014 (ed.), *Nuel Belnap on Indeterminism and Free Action*, Springer, Dordrecht. (Cited on p. 351.)

Mullins, Chris
2005 *The Biogenesis of Cellular Organelles*, Landes Bioscience and Kluwer Academic, Georgetown, Texas and New York. (Cited on p. 276.)

Mumford, Stephen
1994 "Dispositions," *Cogito*, 8(2), pp. 141–46. (Cited on pp. 127, 198.)
1998 *Dispositions*, Oxford University Press, New York. (Cited on pp. 127, 198.)
2004 *Laws in Nature*, Routledge, New York. (Cited on pp. 23, 127, 198.)

Mumford, Stephen and Rani Lill Anjum
2011 *Getting Causes from Powers*, Oxford University Press, New York. (Cited on pp. 23, 127, 198.)

Mumford, Stephen and Matthew Tugby
2013 *Metaphysics and Science*, Oxford University Press, New York. (Cited on p. 338.)

Nagel, Ernest
1961 *The Structure of Science: Problems in the Logic of Scientific Explanation*, Harcourt, Brace & World, New York. (Cited on p. 266.)

Nagel, Thomas
1986 *The View From Nowhere*, Oxford University Press, New York. (Cited on p. 2.)

Noë, Alva and Evan Thompson
2002 (eds.), *Vision and Mind: Selected Readings in the Philosophy of Perception*, MIT Press, Cambridge, Mass. (Cited on p. 35.)

Noonan, Harold W.
2009 "What is a One-Level Criterion of Identity?" *Analysis*, 69(2), pp. 274–7. (Cited on p. 67.)

Normandin, Sebastian and Charles T. Wolfe
2013 (eds.), *Vitalism and the Scientific Image in Post-Enlightenment Life Science, 1800–2010*, Springer, Dordrecht. (Cited on pp. 268, 328, 352.)

Norton, John D.
2007 "Causation as Folk Science," in Price *et al.* [2007], ch. 2, pp. 11–44. (Cited on p. 139.)
2010 "Time Really Passes," *Humana.Mente*, 13, pp. 23–34. (Cited on pp. 219, 253.)

Oaklander, L. Nathan
1991 "A Defence of the New Tenseless Theory of Time," *The Philosophical Quarterly*, 41(162), pp. 26–38. (Cited on p. 215.)

Oaklander, L. Nathan and Quentin Smith
1994 *The New Theory of Time*, Yale University Press, New Haven. (Cited on p. 215.)

Oderberg, David S.
2007 *Real Essentialism*, Routledge, London. (Cited on pp. 23, 44, 60, 62.)
2012 "No Potency Without Actuality," in Tahko [2012], ch. 13, pp. 207–28. (Cited on p. 198.)

Øhrstrøm, Peter
2009 "In Defence of the Thin Red Line: A Case for Ockhamism," *Humana.Mente*, 8, pp. 17–32. (Cited on p. 137.)

Okamoto, Keigo, Sayaka Uchida, Takeru Ito, and Noritaka Mizuno
2007 "Self-Organization of All-Inorganic Dodecatungstophosphate Nanocrystallites," *Journal of the American Chemical Society*, 129(23), pp. 7378–84. (Cited on p. 150.)

Ostwald, Wilhelm
1908 *Grundriss der Naturphilosophie*, 1, Reclam, Leipzig. (Cited on p. 125.)
1929 *Die Pyramide der Wissenschaften*, J.G. Cotta'sche Buchhandlung, Stuttgart und Berlin. (Cited on p. 125.)

Page, Robert B., Meredith A. Boley, Jeramiah J. Smith, Srikrishna Putta, and Stephen R. Voss
2010 "Microarray Analysis of a Salamander Hopeful Monster Reveals Transcriptional Signatures of Paedomorphic Brain Development," *BMC Evolutionary Biology*, 10(1), p. 199. (Cited on p. 303.)

Pasnau, Robert
1997 *Theories of Cognition in the Later Middle Ages*, Cambridge University Press, Cambridge. (Cited on p. 35.)

Pasteur, Louis
1922–39 *Oevres de Pasteur*, Libraires de l'Académie de Médecine, Paris. (Cited on p. 300.)
Paul, L. A.
2006 "Coincidence as Overlap," *Noûs*, 40(4), pp. 623–59. (Cited on p. 24.)
2010 "The Puzzles of Material Constitution," *Philosophy Compass*, 5(7), pp. 579–90. (Cited on p. 24.)
Peacocke, Christopher
1976 "What Is a Logical Constant?" *Journal of Philosophy*, 73, pp. 221–40. (Cited on p. 100.)
1987 "Understanding Logical Constants: A Realist's Account," *Proceedings of the British Academy*, 73, pp. 153–200. (Cited on p. 102.)
Pearl, Judea
2000 *Causality: Models, Reasoning, and Inference*, New York, Cambridge University Press. (Cited on p. 187.)
Peterson, Daniel and Michael Silberstein
2010 "Relativity of Simultaneity and Eternalism: In Defense of Blockworld," in Petkov [2010], pp. 209–38. (Cited on p. 252.)
Petkov, Vesselin
2010 *Space, Time, and Spacetime: Physical and Philosophical Implications of Minkowski's Unification of Space and Time*, Springer, Dordrecht. (Cited on pp. 346, 354.)
Pickavance, Timothy
2009 "In Defence of 'Partially Clad' Bare Particulars," *Australasian Journal of Philosophy*, 87(1), pp. 155–58. (Cited on p. 79.)
Plantinga, Alvin
1974 *The Nature of Necessity*, Oxford University Press. (Cited on p. 62.)
1993 *Warrant and Proper Function*, Oxford University Press, New York. (Cited on pp. 289, 294.)
Pooley, Oliver
2013 "Relativity, the Open Future, and the Passage of Time," *Proceedings of the Aristotelian Society*, 113(3), pp. 321–63. (Cited on pp. 221, 229, 249, 251, 253.)
Post, H.
1963 "Individuality and Physics," *The Listener*, 70, pp. 534–7. (Cited on p. 273.)
Powell, Alexander and John Dupré
2009 "From Molecules to Systems: The Importance of Looking Both Ways," *Studies in History and Philosophy of Biological and Biomedical Sciences*, 40(1), pp. 54–64. (Cited on p. 271.)
Prawitz, Dag
1985 "Remarks on Some Approaches to the Concept of Logical Consequence," *Synthese*, 62(2), pp. 153–71. (Cited on p. 102.)
Price, Huw and Richard Corry
2007 (eds.), *Causation, Physics, and the Constitution of Reality: Russell's Republic Revisited*, Oxford University Press, New York. (Cited on pp. 341, 345.)
Priest, Graham
2006 *In Contradiction*, Oxford University Press, Oxford. (Cited on p. 217.)
Prior, Arthur N.
1957 *Time and Modality*, Oxford University Press, Oxford. (Cited on pp. 216, 218, 234.)
1960 "The Runabout Inference-Ticket," *Analysis*, 21(2), pp. 38–9. (Cited on p. 102.)
1967 *Past, Present and Future*, Oxford University Press, Oxford. (Cited on pp. 137, 142, 214, 216, 218.)
1968a "Changes in Events and Changes in Things," in Prior [1968b], ch. 1, pp. 1–14. (Cited on pp. 226, 229.)
1968b *Papers on Time and Tense*, Oxford University Press, Oxford. (Cited on pp. 218, 346.)

Prior, Arthur N.
1976 *Papers in Logic and Ethics*, University of Massachusetts Press, Amherst. (Cited on pp. 216, 218.)

Prior, Arthur N. and Kit Fine
1977 *Worlds, Times and Selves*, University of Massachusetts Press, Amherst. (Cited on pp. 218, 230, 234, 236.)

Putnam, Hilary
1975 "The Meaning of 'Meaning'," *Minnesota Studies in the Philosophy of Science*, 7, pp. 131–93. (Cited on pp. 63, 118, 280.)
1977 "Realism and Reason," *Proceedings and Addresses of the American Philosophical Association*, 50, pp. 483–98. (Cited on p. 30.)
1981 *Reason, Truth and History*, Cambridge University Press, Cambridge, Mass. (Cited on p. 32.)
1983 *Philosophical Papers: Volume 3, Realism and Reason*, Cambridge University Press, Cambridge, Mass. (Cited on p. 32.)
1991 *Representation and Reality*, The MIT Press, Cambridge, Mass. (Cited on p. 52.)
1992 *Realism with a Human Face*, Harvard University Press, Cambridge, Mass. (Cited on pp. 17, 52.)
1999 *The Threefold Cord*, Columbia University Press, New York. (Cited on pp. 17, 52.)

Putnam, Hilary and James Conant
1994 *Words and Life*, Harvard University Press, Cambridge, Mass. (Cited on p. 52.)

Quine, Willard V.
1943 "Notes on Existence and Necessity," *Journal of Philosophy*, 40(5), pp. 113–27. (Cited on pp. 114, 167.)
1947 "The Problem of Interpreting Modal Logic," *Journal of Symbolic Logic*, 12(2), pp. 43–8. (Cited on pp. 114, 167.)
1948 "On What There Is," *Review of Metaphysics*, 2(5), pp. 21–36. (Cited on pp. 114, 167.)
1950 *Methods of Logic*, Holt, Reinhart and Winston, New York. (Cited on p. 90.)
1951 "Two Dogmas of Empiricism," *The Philosophical Review*, 60(1), pp. 20–43. (Cited on p. 116.)
1953a *From a Logical Point of View*, Harvard University Press, Cambridge, Mass. (Cited on p. 347.)
1953b "Reference and Modality," in Quine [1953a]. (Cited on pp. 114, 167.)
1953c "Three Grades of Modal Involvement," in *Proceedings of the XIth International Congress of Philosophy*, North-Holland Publishing Co., Amsterdam. (Cited on pp. 114, 167.)
1957 "Speaking of Objects," *Proceedings and Addresses of the American Philosophical Association*, 31, pp. 5–22. (Cited on p. 71.)
1960 *Word and Object*, The MIT Press, Cambridge, Mass. (Cited on pp. 77, 81, 197.)
1964 "Ontological Reduction and the World of Numbers," *The Journal of Philosophy*, 61(7), pp. 209–16. (Cited on pp. 25, 162.)
1970 *Philosophy of Logic*, Harvard University Press, Cambridge, Mass. (Cited on pp. 100, 215.)
1976a "Whither Physical Objects?" In R. S. Cohen *et al.* [1976], pp. 497–504. (Cited on p. 162.)
1976b "Worlds Away," *Journal of Philosophy*, 73(22), pp. 859–63. (Cited on pp. 114, 167.)
1985 "Events and Reification," in Davidson [2006], ch. 11, pp. 162–71. (Cited on pp. 77, 163.)
1991 "Two Dogmas in Retrospect," *Canadian Journal of Philosophy*, 21(3), pp. 265–74. (Cited on pp. 114, 167.)

Ramsey, Frank P.
1928 "Universals of Law and Fact," in Ramsey [1978], ch. 6A, pp. 128–32. (Cited on p. 172.)
1929 "General Propositions and Causality," in Ramsey [1978], ch. 6B, pp. 133–51. (Cited on p. 172.)
1978 *Foundations: Essays in Philosophy, Logic, Mathematics and Economics*, Revised edition, Routledge & Kegan Paul, London and Henley. (Cited on p. 347.)

Rand, Patricia
2001 *CliffsQuickReview Plant Biology*, IDG Books Worldwide, New York. (Cited on p. 149.)

Ray, Greg
1996 "Logical Consequence: A Defense of Tarski," *Journal of Philosophical Logic*, 25(6), pp. 617–77. (Cited on pp. 98, 105.)

Rea, Michael C.
1995 "The Problem of Material Constitution," *The Philosophical Review*, 104(4), pp. 525–52. (Cited on p. 24.)
2002 *World without Design: The Ontological Consequences of Naturalism*, Oxford University Press. (Cited on p. 294.)

Reichenbach, Hans
1947 *Elements of Symbolic Logic*, Macmillan, New York. (Cited on p. 215.)

Reicher, Maria E. and Johan C. Marek
2005 *Experience and Analysis. The Proceedings of the 27th International Wittgenstein Symposium*, öbv&hpt, Vienna. (Cited on p. 350.)

Rescher, Nicholas
1996 *Process Metaphysics: An Introduction to Process Philosophy*, State University of New York Press, Albany. (Cited on p. 137.)
2000 *Process Philosophy: A Survey of Basic Issues*, University of Pittsburgh Press, Pittsburgh. (Cited on p. 137.)

Reutlinger, Alexander
2012 "Getting Rid of Interventions," *Studies in History and Philosophy of Science Part C*, 43(4), pp. 787–95. (Cited on p. 188.)

Rey, Georges
2012 "The Analytic/Synthetic Distinction," in *The Stanford Encyclopedia of Philosophy*, ed. by Edward N. Zalta, Summer 2012 version. (Cited on p. 116.)

Reydon, Thomas A. C.
2008 "Species in Three and Four Dimensions," *Synthese*, 164(2), pp. 161–84. (Cited on p. 304.)

Richards, Richard A.
2010 *The Species Problem: A Philosophical Analysis*, Cambridge University Press, Cambridge. (Cited on p. 301.)

Rödl, Sebastian
2007 *Self-Consciousness*, Harvard University Press, Cambridge, Mass. (Cited on p. 17.)
2012/2005 *Categories of the Temporal: an Inquiry Into the Forms of the Finite Intellect*, Translated from the German by Sibylle Salewski, Harvard University Press, Cambridge, Mass. (Cited on pp. 17, 138, 207, 232, 236, 237, 241, 242.)

Rodriguez-Pereyra, Gonzalo
2002 *Resemblance Nominalism: A Solution to the Problem of Universals*, Oxford University Press, New York. (Cited on pp. 18, 25, 27.)

Rogozin, Igor B., Liran Carmel, Miklos Csuros, and Eugene V. Koonin
2012 "Origin and Evolution of Spliceosomal Introns," *Biology Direct*, 7(11), pp. 1–28. (Cited on p. 276.)

Roll-Hansen, Nils
1979 "Experimental Method and Spontaneous Generation: The Controversy between Pasteur and Pouchet, 1859-64," *Journal of the History of Medicine and Allied Sciences*, 34(3). (Cited on p. 300.)

Rorty, Richard
1989 *Contingency, Irony, and Solidarity*, Cambridge University Press, Cambridge, Mass. (Cited on p. 36.)

Rosenberg, Alexander
1985 *The Structure of Biological Science*, The University of Chicago Press, Chicago. (Cited on p. 266.)
1994 *Instrumental Biology or the Disunity of Science*, University of Chicago Press, Chicago. (Cited on p. 266.)
2006 *Darwinian Reductionism, or: How to Stop Worrying and Love Molecular Biology*, University of Chicago Press, Chicago. (Cited on p. 266.)

Rumberg, Antje
n.d. *Transitions Towards a Semantics for Real Possibility*, Manuscript, PhD thesis. (Cited on pp. 142, 249.)
Rumberg, Antje and Thomas Müller
2013 "Transitions Towards a New Account of Future Contingents," Unpublished manuscript. (Cited on pp. 142, 249.)
Russell, Bertrand
1903 *The Principles of Mathematics*, Cambridge University Press, Cambridge. (Cited on p. 163.)
1915 "On the Experience of Time," *The Monist*, 25(2), pp. 212–33. (Cited on p. 220.)
1919 "The Philosophy of Logical Atomism," *The Monist*, 29(1), pp. 32–63. (Cited on p. 215.)
1941 *An Inquiry into Meaning and Truth*, George Allen & Unwin Ltd., London. (Cited on p. 215.)
1953 "On the Notion of Cause, with Application to the Free-Will Problem," in Feigl *et al.* [1953], pp. 387–418. (Cited on p. 250.)
Rychter, Pablo
2008 "Perdurance, Endurance, and 'Having a Property *Atemporally*'," *Metaphysica*, 9(2), pp. 159–71. (Cited on pp. 207, 210.)
Ryder, Dan, Justine Kingsbury, and Kenneth Williford
2013 (eds.), *Millikan and Her Critics*, Wiley-Blackwell, Chichester, West Sussex. (Cited on pp. 279, 334, 344.)
Ryle, Gilbert
1949 *The Concept of Mind*, Hutchinson, London. (Cited on p. 197.)
1954 *Dilemmas*, Cambridge University Press, Cambridge. (Cited on p. 100.)
Sainsbury, Mark
2001 *Logical Forms: An Introduction to Philosophical Logic*, 2nd edition, Blackwell Publishing, Oxford. (Cited on p. 102.)
Sapienza, Carmen
2010 "Selection Does Operate Primarily on Genes: In Defense of the Gene as the Unit of Selection," in Ayala *et al.* [2010], ch. 7, pp. 127–40. (Cited on p. 298.)
Saunders, Simon
2002 "How Relativity Contradicts Presentism," in Callender [2002], pp. 277–92. (Cited on p. 252.)
Scaltsas, Theodore
1994 *Substances and Universals in Aristotle's Metaphysics*, Cornell University Press, Ithaca. (Cited on p. 82.)
Schaffer, Jonathan
2014 *The Metaphysics of Causation*, ed. by Edward N. Zalta, http://plato.stanford.edu/archives/sum2014/entries/causation-metaphysics/. (Cited on p. 184.)
Schaffner, Kenneth F.
1969 "The Watson-Crick Model and Reductionism," *British Journal for the Philosophy of Science*, 20(4), pp. 235–48. (Cited on p. 266.)
Schlick, Moritz
1949 *Philosophy of Nature*, Philosophical Library, New York. (Cited on p. 259.)
Schrenk, Markus
2006 "A Theory for Special Science Laws," in *Selected Papers Contributed to the Sections of GAP.6*, mentis. (Cited on p. 174.)
2011 "Interfering with Nomological Necessity," *The Philosophical Quarterly*, 61(244), pp. 577–97. (Cited on p. 196.)
Sellars, Wilfrid S.
1958 "Counterfactuals, Dispositions and the Causal Modalities," in *Minnesota Studies in the Philosophy of Science*, ed. by G. Maxwell H. Feigl M. Scriven, University of Minnesota Press, Minneapolis, vol. 2, pp. 225–308. (Cited on p. 197.)

Sellars, Wilfrid S.

1962 "Philosophy and the Scientific Image of Man," in Colodny [1962], pp. 35–68. (Cited on pp. 33, 34.)

1981a "Foundations for a Metaphysics of Pure Process, I: The Lever of Archimedes," *The Monist*, 64(1), pp. 3–36. (Cited on p. 137.)

1981b "Foundations for a Metaphysics of Pure Process, II: Naturalism and Process," *The Monist*, 64(1), pp. 37–65. (Cited on p. 137.)

1981c "Foundations for a Metaphysics of Pure Process, III: Is Consciousness Physical?" *The Monist*, 64(1), pp. 66–90. (Cited on p. 137.)

Sher, Gila Y.

1991 *The Bounds of Logic: a Generalized Viewpoint*, MIT Press, Cambridge, Mass. (Cited on p. 101.)

1996 "Did Tarski Commit "Tarski's Fallacy"?" *The Journal of Symbolic Logic*, 61(2), pp. 653–86. (Cited on p. 101.)

2003 "A Characterization of Logical Constants *Is* Possible," *Theoria: Revista de Teoria, Historia y Fundamentos de la Ciencia*, 18, pp. 189–97. (Cited on p. 101.)

Shoemaker, Sydney

1969 "Time Without Change," *Journal of Philosophy*, 66(12), pp. 363–81. (Cited on p. 247.)

Sidelle, Alan

1989 *Necessity, Essence, and Individuation: A Defense of Conventionalism*, Cornell University Press. (Cited on pp. 31, 60.)

1998 "A Sweater Unraveled: Following One Thread of Thought for Avoiding Coincident Entities," *Noûs*, 32(4), pp. 423–48. (Cited on pp. 24, 31, 60.)

2010 "Modality and Objects," *The Philosophical Quarterly*, 60(238), pp. 109–25. (Cited on p. 60.)

Sider, Theodore

2001 *Four-Dimensionalism: An Ontology of Persistence and Time*, Oxford University Press, New York. (Cited on pp. 22, 24, 127, 162, 163, 168, 205, 207–209.)

2006a ""Bare Particulars"," *Philosophical Perspectives*, 20(1), pp. 387–97. (Cited on p. 79.)

2006b "Beyond the Humphrey Objection," `http://tedsider.org/papers/counterpart_theory.pdf`, (visited on 03/31/2014). (Cited on p. 167.)

2009 "Ontological Realism," in Chalmers *et al.* [2009], ch. 13, pp. 384–423. (Cited on p. 18.)

2012 *Writing the Book of the World*, Oxford University Press, New York. (Cited on pp. 18, 22, 25, 127, 179, 209.)

Sider, Theodore, John Hawthorne, and Dean W. Zimmerman

2008 (eds.), *Contemporary Debates in Metaphysics*, Wiley-Blackwell, Malden, Mass. (Cited on pp. 333, 354.)

Simons, Peter

1987 *Parts: A Study in Ontology*, Oxford University Press, New York. (Cited on pp. 77, 163.)

1994 "Particulars in Particular Clothing: Three Trope Theories of Substance," *Philosophy and Phenomenological Research*, 54(3), pp. 553–75. (Cited on p. 70.)

Skow, Bradford

2009 "Relativity and the Moving Spotlight," *Journal of Philosophy*, 106(12), pp. 666–78. (Cited on pp. 220, 221.)

Smart, J. J. C.

1959 "Can Biology Be an Exact Science?" *Synthese*, 11(4), pp. 359–68. (Cited on p. 276.)

1963 *Philosophy and Scientific Realism*, Routledge & Kegan Paul, London. (Cited on pp. 161, 215, 248.)

Smith, Barry

2005 "Against Fantology," in Reicher *et al.* [2005], pp. 153–70. (Cited on p. 116.)

Smith, Quentin

1993 *Language and Time*, Oxford University Press, New York. (Cited on p. 220.)

2002 "Time and Degrees of Existence: A Theory of "Degree Presentism"," in Callender [2002], pp. 119–36. (Cited on p. 220.)

Sosa, Ernest
1993 "Putnam's Pragmatic Realism," *The Journal of Philosophy*, 90(12), pp. 605–26. (Cited on p. 174.)
1999 "Existential Relativity," *Midwest Studies in Philosophy*, 23(1), pp. 132–43. (Cited on pp. 32, 174.)

Sprevak, Mark D.
2008 "Kripke's Paradox and the Church–Turing Thesis," *Synthese*, 160(2), pp. 285–95. (Cited on p. 283.)

Stanford, P. Kyle
2012 "The Eyes Don't Have It: Fracturing the Scientific and Manifest Images," *Humana.Mente*, 21, pp. 19–44. (Cited on p. 34.)

Steiner, Rudolf
1886 *Grundlinien einer Erkenntnistheorie der Goetheschen Weltanschauung mit besonderer Rücksicht auf Schiller*, 6. Auflage, 1960, Rudolf Steiner Verlag, Dornach, Schweiz. (Cited on pp. 18, 125.)
1894 *Die Philosophie der Freiheit: Grundzüge einer modernen Weltanschauung*. Verlag Emil Felber, Berlin. (Cited on pp. 18, 125.)

Strawson, Peter F.
1966 *The Bounds of Sense: An Essay on Kant's Critique of Pure Reason*, Methuen, London. (Cited on p. 30.)

Strobach, Niko
2014 "In Retrospect: Can BST Models be Reinterpreted for What Decisions, Speciation Events and Ontogeny Might Have in Common?" In Müller [2014], pp. 259–78. (Cited on p. 307.)

Stroud, Barry
1984 *The Significance of Philosophical Scepticism*, Oxford University Press, New York. (Cited on p. 31.)
2000 *The Quest for Reality: Subjectivism & the Metaphysics of Colour*, Oxford University Press, New York. (Cited on pp. 2, 47–49, 52, 54.)
2011 *Engagement and Metaphysical Dissatisfaction: Modality and Value*, Oxford University Press, New York. (Cited on pp. 2, 47, 50–55, 91, 121, 184.)

Swartz, Norman
1985 *The Concept of Physical Law*, Cambridge University Press, Cambridge. (Cited on p. 171.)

Swoyer, Chris
1982 "The Nature of Natural Laws," *Australasian Journal of Philosophy*, 60(3), pp. 203–23. (Cited on p. 198.)

Tahko, Tuomas E.
2012 (ed.), *Contemporary Aristotelian Metaphysics*, Cambridge University Press, Cambridge. (Cited on pp. 23, 342, 345.)

Tarski, Alfred
1935 "Der Wahrheitsbegriff in den formalisierten Sprachen," *Studia Philosophica*, I, pp. 261–405. (Cited on p. 98.)
1936 "On the Concept of Logical Consequence," in Tarski [1956], pp. 409–20. (Cited on pp. 92, 93, 97, 106.)
1956 *Logic, Semantics, Metamathematics*, Clarendon Press, Oxford. (Cited on p. 351.)
1986 "What Are Logical Notions?" *History and Philosophy of Logic*, 7, Transcript of a 1966 talk, ed. J. Corcoran., pp. 143–54. (Cited on p. 101.)

Taylor, Barry
1993 "On Natural Properties in Metaphysics," *Mind*, 102(405), pp. 81–100. (Cited on pp. 173, 174.)

Theissen, Günther
2005 "The Proper Place of Hopeful Monsters in Evolutionary Biology," *Theory in Biosciences*, 124, pp. 349–69. (Cited on p. 303.)

Thomasson, Amie L.
2007 *Ordinary Objects*, Oxford University Press, New York. (Cited on p. 60.)
Thompson, Michael
2008 *Life and Action*, Harvard University Press, Cambridge, Mass. (Cited on pp. 17, 18, 43, 114, 121, 125, 128, 150, 259–264, 282, 325, 366, 371, 378.)
Thomson, Judith Jarvis
1983 "Parthood and Identity Across Time," *The Journal of Philosophy*, 80(4), pp. 201–20. (Cited on p. 24.)
1998 "The Statue and the Clay," *Noûs*, 32(2), pp. 149–73. (Cited on p. 24.)
Tooley, Michael
1977 "The Nature of Laws," *Canadian Journal of Philosophy*, 7(4), pp. 667–98. (Cited on p. 189.)
Torrengo, Giuliano
2013 "The Grounding Problem and Presentist Explanations," *Synthese*, 190(12), pp. 2047–63. (Cited on p. 226.)
Trettin, Käthe
2005 (ed.), *Substanz: Neue Überlegungen zu einer klassischen Kategorie des Seienden*, Klostermann, Frankfurt/Main. (Cited on p. 353.)
Turner, J. Scott
2013 "Homeostasis and the Forgotten Vitalist Roots of Adaptation," in Normandin *et al.* [2013], ch. 11, pp. 271–91. (Cited on p. 268.)
Unger, Peter
1979a "I Do Not Exist," in Unger [2006], ch. 2, pp. 36–52. (Cited on p. 20.)
1979b "There Are No Ordinary Things," in Unger [2006], ch. 1, pp. 3–35. (Cited on p. 20.)
1979c "Why There Are No People," in Unger [2006], ch. 3, pp. 53–109. (Cited on p. 20.)
1980 "The Problem of the Many," in Unger [2006], ch. 4, pp. 113–82. (Cited on p. 20.)
2006 *Philosophical Papers: Volume Two*, Oxford University Press, New York. (Cited on p. 352.)
Van Benthem, Johan
1989 "Logical Constants Across Varying Types," *Notre Dame Journal of Formal Logic*, 30, pp. 315–42. (Cited on p. 101.)
Van Cleve, James
1985 "Why a Set Contains its Members Essentially," *Noûs*, 19(4), pp. 585–602. (Cited on pp. 66, 80.)
Van Inwagen, Peter
1990 "Four-Dimensional Objects," *Noûs*, 24(2), pp. 245–55. (Cited on p. 207.)
1998 "Meta-Ontology," *Erkenntnis*, 48(2), pp. 233–50. (Cited on p. 75.)
Van Inwagen, Peter and Dean W. Zimmerman
1998 *Metaphysics: the Big Questions*, Blackwell Publishing, Malden, Mass. (Cited on p. 354.)
Von Wright, Georg Henrik
1971 *Explanation and Understanding*, Cornell University Press, Ithaca, New York. (Cited on p. 187.)
Ward, Barry
2002 "Humeanism Without Humean Supervenience: A Projectivist Account of Laws and Possibilities," *Philosophical Studies*, 107(3), pp. 191–208. (Cited on p. 178.)
Warmbröd, Ken
1999 "Logical Constants," *Mind*, 108(431), pp. 503–38. (Cited on p. 105.)
Weinert, Friedel
2010 "Relativistic Thermodynamics and the Passage of Time," *Humana.Mente*, 13, pp. 175–91. (Cited on p. 253.)
Whitehead, Alfred N.
1920 *The Concept of Nature*, Cambridge University Press, Cambridge. (Cited on p. 137.)

Whitehead, Alfred N.
1929 *Process and Reality*, Cambridge University Press, Cambridge. (Cited on p. 137.)
Whitford, David
2005 *Proteins: Structure and Function*, Wiley, Chichester. (Cited on p. 275.)
Whittle, Ann
2008 "A Functionalist Theory of Properties," *Philosophy and Phenomenological Research*, 77(1), pp. 59–82. (Cited on p. 198.)
Wiggins, David
1968 "On Being in the Same Place at the Same Time," *The Philosophical Review*, 77(1), pp. 90–5. (Cited on p. 24.)
1995 "Substance," in Trettin [2005], pp. 105–45. (Cited on pp. 23, 60.)
2001 *Sameness and Substance Renewed*, 2nd edition, Cambridge University Press, Cambridge. (Cited on pp. 17, 23, 24, 44, 60, 62, 66, 136, 168, 208, 231.)
Wilkins, John S.
2006 *A List of 26 Species "Concepts"*, http://scienceblogs.com/evolvingthoughts/2006/10/01/a-list-of-26-species-concepts/ (visited on 12/09/2013). (Cited on p. 305.)
2009 *Species: A History of the Idea*, University of California Press, Berkeley. (Cited on pp. 303, 304.)
2010 "What is a Species? Essences and Generation," *Theory in Biosciences*, 129, pp. 141–8. (Cited on pp. 302–305.)
Williams, Bernard
1978 *Descartes*, Penguin Books, London. (Cited on pp. 2, 18.)
Williams, Donald C.
1951 "The Myth of Passage," *Journal of Philosophy*, 48(15), pp. 457–72. (Cited on pp. 215, 221, 224.)
Williamson, Timothy
1991 "Fregean Directions," *Analysis*, 51(4), pp. 194–95. (Cited on p. 66.)
1994 *Vagueness*, Routledge, New York. (Cited on p. 307.)
1999 "Logic and Existence," *Aristotelian Society Supplementary Volume*, 73(1), pp. 181–203. (Cited on pp. 114, 167.)
2010 "Necessitism, Contingentism, and Plural Quantification," *Mind*, 119(475), pp. 657–748. (Cited on pp. 114, 167.)
2013 *Modal Logic as Metaphysics*, Oxford University Press, Oxford. (Cited on pp. 24, 95, 106, 114, 127, 167, 211.)
Winsor, Mary P.
2006 "The Creation of the Essentialism Story: An Exercise in Metahistory," *History and Philosophy of the Life Sciences*, 28(2), pp. 149–74. (Cited on p. 302.)
Wittgenstein, Ludwig
1921 *Tractatus Logico-Philosophicus*, Suhrkamp Verlag, Frankfurt am Main. (Cited on pp. 20, 201.)
1953 *Philosophical Investigations*, Blackwell Publishing, Oxford. (Cited on pp. 20, 52.)
Woodger, Joseph H.
1929 *Biological Principles: A Critical Study*, Routledge & Kegan Paul, London. (Cited on p. 269.)
Woodward, James
1997 "Explanation, Invariance, and Intervention," *Philosophy of Science*, 64(4), pp. 26–41. (Cited on p. 187.)
2003 *Making Things Happen: A Theory of Causal Explanation*, Oxford University Press, New York. (Cited on pp. 187, 188.)
Woodward, James and Christopher Hitchcock
2003 "Explanatory Generalizations, Part I: A Counterfactual Account," *Noûs*, 37(1), pp. 1–24. (Cited on p. 187.)

Wüthrich, Christian

2010 "No Presentism in Quantum Gravity," in Petkov [2010], pp. 257–79. (Cited on p. 252.)

2013 "The Fate of Presentism in Modern Physics," in Ciuni *et al.* [2013], pp. 91–132. (Cited on p. 254.)

Yablo, Stephen

1987 "Identity, Essence, and Indiscernibility," *The Journal of Philosophy*, 84(6), pp. 293–314. (Cited on p. 62.)

Zimmerman, Dean W.

1998 "Temporary Intrinsics and Presentism," in van Inwagen *et al.* [1998], ch. 25, pp. 206–19. (Cited on p. 212.)

2004 (ed.), *Oxford Studies in Metaphysics, Vol. 1*, Clarendon Press, Oxford. (Cited on pp. 331, 342.)

2005 "The A-Theory of Time, The B-Theory of Time, and 'Taking Tense Seriously'," *dialectica*, 59(4), pp. 401–57. (Cited on pp. 214–216, 254.)

2006 (ed.), *Oxford Studies in Metaphysics, Vol. 2*, Oxford University Press, New York. (Cited on pp. 335, 343.)

2008 "The Priviledged Present: Defending an "A-theory" of Time," in Sider *et al.* [2008], ch. 5.1, pp. 211–25. (Cited on pp. 216, 254.)

2011 "Presentism and the Space-Time Manifold," in Callender [2011b], ch. 7, pp. 163–244. (Cited on p. 254.)

Samenvatting in het Nederlands

Dit proefschrift gaat over metafysica, en het gaat daarover op twee verschillende manieren. Ten eerste behandel ik de vraag wat metafysica nu eigenlijk is, en van welk uitgangspunt metafysica moet uitgaan. Dat is het onderwerp van deel I; daarin heb ik het *over* metafysica. Ten tweede ontwikkel ik in deze dissertatie een metafysica, op basis van het uitgangspunt dat in het eerste deel geformuleerd en verdedigd is. Dat is het onderwerp van deel II; daarin *bedrijf* ik dus metafysica.[392]

Dat gezegd hebbende, ligt de vraag voor de hand: wat is dat dan precies, metafysica? Als opmaat voor deze samenvatting van mijn onderzoeksresultaten wijd ik graag een paar woorden aan deze vraag. Een opmerking vooraf: mijn onderzoek orienteert zich aan de hedendaagse metafysische discussie die zich binnen de kaders van de *analytische wijsbegeerte* afspeelt. Dat is een brede (en moeilijk te definiëren) stroming in de filosofie die zijn wortels heeft in het werk van logici en wetenschapsfilosofen uit het begin van de twintigste eeuw.

Aristoteles' *Metafysica* opent met de beroemde constatering: 'Alle mensen streven van nature naar inzicht' [Aristoteles 1998, p. 4, 980a21; vertaling JM]. En dat klopt: of het nu gaat om onze dagelijkse beslommeringen, of om het systematische, academische onderzoek waar onze wetenschappers hun leven aan wijden: we vinden het belangrijk om te begrijpen waar we zoal mee te maken hebben, en om ons leven overeenkomstig dat verworven inzicht vorm te geven.

Hoe doen we dat precies, tot inzicht komen? Hoe vinden we onze weg in de werkelijkheid? Het antwoord is bedrieglijk eenvoudig: we nemen waar, en we denken na. We bezien hoe onze buurman reageert op ons goedgemutste 'Goedemorgen!', en leren zo om zijn stemming te herkennen—zodat we weten wanneer we hem voor

[392] In deze Nederlandse samenvatting beperk ik mij tot het wezenlijke: een korte en begrijpelijke weergave van de inhoud van dit proefschrift. Alleen waar dat echt nodig is zal ik, liefst in voetnoten, verwijzen naar relevante bronnen en naar passages uit de hoofdtekst.

hulp in de tuin kunnen vragen, en wanneer beter niet. We merken op wat er met de bacteriën in ons petrischaaltje gebeurt nadat we er penicilline aan toegevoegd hebben, en besluiten dat dat goedje wel eens bruikbaar zou kunnen zijn voor medische doeleinden. We analyseren een eigenaardig patroon in het spectrum van licht dat uit een verre uithoek van het heelal afkomstig is, en speculeren dat in die uithoek waterstof te vinden is.

In deze en dergelijke situaties gebruiken we ons denkvermogen om ons inzicht in wat we voor ons hebben te vergroten: we observeren, en we denken na. Als we alleen de waarneming hadden, dan zou er slechts een aaneenschakeling van ervaringen zijn, zonder enige samenhang: inzicht is pas mogelijk wanneer we denkend op onze waarnemingen ingaan. Als we onder *begrippen* of *concepten* de elementen van dat denken verstaan, dan kunnen we dus zeggen dat onze kennis, ons inzicht in de werkelijkheid *begripsmatig*, *conceptueel* is.

Terugkomend op onze voorbeelden zien we dat we ons met de genoemde stukjes werkelijkheid uiteenzetten door begrippen te vatten—begrippen zoals, naast vele anderen, *mens*, *karakter* en *stemming*; *bacterie* en *penicilline*; en *licht*, *spectrum* en *waterstof*. Overdenk die voorbeelden eens, en bekijk hoe de relevante begrippen zodoende denkend met elkaar verbonden worden: *dat* is wat ik bedoel met 'begripsmatig' of 'conceptueel'.

Zoals Aristoteles schreef streven wij mensen dus naar inzicht, en inzicht is, in de geschetste zin, conceptueel. Metafysica is nu de discipline die zich dit streven van ons in zijn meest algemene en overkoepelende vorm tot opgave maakt. Net als alle filosofie is ook metafysica een vorm van onderzoek waarbij de nadruk ligt op denken, niet op waarnemen. Daarmee is niet gezegd dat metafysica onafhankelijk is van de waarneming, of zich niets hoeft aan te trekken van de zogeheten 'empirische wetenschap'—het is immers een vorm van onderzoek dat uitgevoerd wordt door *ons*, wezens behept met zowel zintuiglijke als verstandelijke vermogens. Ter vergelijking: we kunnen op een goede dag reflecteren op onze ontmoetingen met bovengenoemde buurman, en zodoende tot overwegingen komen die over heel algemene zaken gaan, zoals het menselijk karakter, sociale interactie, vriendschap, en dergelijken. Zulk soort bezinning speelt zich onafhankelijk van de waarneming af, en is dus in zekere zin niet-empirisch, een puur conceptuele bezigheid. Op dezelfde manier is metafysica niet-empirisch en puur conceptueel, zonder daarmee los te staan van de waarneembare werkelijkheid. Integendeel: metafysisch onderzoek begint natuurlijk ook met vragen die we ons bij onze ervaringen stellen, en wordt voortdurend gevoed door de inzichten die we in onze omgang met de werkelijkheid eigen maken.

In de metafysica stellen we vragen waar onze alledaagse en wetenschappelijke beschouwingen en onderzoekingen vaak niet expliciet aan raken. Een dergelijke vraag is bijvoorbeeld de volgende: wat is de relatie tussen ons inzicht in de werkelijkheid, dat in de geschetste zin conceptueel is, en de werkelijkheid zelf? Die vraag ligt impliciet aan alle onderzoek, in het groot en in het klein, ten grondslag: we gaan er stilzwijgend vanuit dat ons onderzoek, dat zich in het begripsmatige voltrekt, iets met de werkelijkheid te maken heeft—maar wat precies? Deze vraag vormt het uitgangspunt van mijn onderzoek; zowel in deel I, waarin ik onderzoek wat de juiste uitgangspunten voor metafysica zijn, alsook in deel II, waarin ik op basis van de verworven uitgangspunten mijn eigen metafysische resultaten schetsmatig ontwikkel.

De onderzoeksvraag die in deze dissertatie centraal staat luidt dus: *wat is de relatie tussen begrip en werkelijkheid?* Aan de ene kant is dit een vraag *over* metafysica: wat is de verhouding tussen het begrippenkader waar de metafysica als onderzoeksvorm naar streeft en de werkelijkheid zelf? Hoe kan metafysica, als puur conceptuele vorm van onderzoek, ons nader tot de werkelijkheid leiden? Op deze eerste wijze begrepen is het dus een *meta*-metafysische vraag, een vraag die naar het wezen en de methode van metafysica als onderzoeksvorm vraagt. Deze metametafysische vraag staat centraal in Deel I.

Aan de andere kant is de genoemde vraag echter ook een vraag *van* de metafysica: hoe verhouden zich de begrippen waarmee wij de werkelijkheid begrijpen tot de werkelijkheid zelf? Op deze tweede wijze begrepen gaat het om een metafysische vraag, een vraag over begrip en werkelijkheid waar de metafysica antwoord op wil geven. Deze metafysische vraag staat centraal in Deel II.

Het antwoord op onze vraag, in beide hoedanigheden, kan kernachtig samengevat worden in één en dezelfde stelling: *begrippen zijn essenties*. Met de 'essentie' van iets bedoel ik hier het wezen ervan: de stelling zegt dus dat het begrip waarmee wij een of ander stukje werkelijkheid begrijpen identiek is met de essentie van dat stukje werkelijkheid, identiek is met de kern van de zaak.

Deze stelling beantwoordt onze onderzoeksvraag in haar metametafysische zin als volgt: begrippen zijn essenties, dus een puur conceptuele onderzoeksvorm zoals metafysica gaat heel direct over de werkelijkheid *juist omdat* het puur conceptueel, begripsmatig is. Negatief uitgedrukt: de vooronderstelling dat er nog een of andere theorie nodig is die ons vertelt hoe onze begrippen zich tot een niet-begripsmatige werkelijkheid zouden verhouden is een misvatting. Deze gedachte ontwikkel ik in meer detail hieronder, in mijn samenvatting van Deel I.

Dezelfde stelling biedt ook het uitgangspunt voor een antwoord op onze onderzoeksvraag in haar metafysische zin: begrippen zijn essenties, en daarom moet de metafysica ernaar streven om de meest algemene begrippen waarmee wij de werkelijkheid begrijpen zo helder en volledig mogelijk te doordringen en uiteen te zetten. Negatief uitgedrukt: de vooronderstelling dat we een minimale metafysische theorie moeten ontwikkelen die als basis dient voor een reductie van alle verdere verschijnselen is een misvatting (wat ik onder 'reductie' versta leg ik hieronder kort uit, in mijn samenvatting van hoofdstuk 2). Het resultaat is een metafysische positie die door en door anti-reductionistisch is. Deze gedachten, en de genoemde metafysische positie, ontwikkel ik in meer detail hieronder, in mijn samenvatting van Deel II.

Deel I: Metametafysica

Metametafysica is nadenken over de mogelijkheid en methoden van metafysisch onderzoek—net als meta-ethiek nadenken over de mogelijkheid en methoden van ethisch onderzoek inhoudt. Het doel van dit eerste, metametafysische deel van mijn proefschrift is tweeledig: enerzijds onderneem ik een kritische zoektocht naar het basale, metametafysische beeld dat ten grondslag ligt aan wat er tegenwoordig binnen de metafysische discussie zoal beweerd en beargumenteerd wordt, en anderzijds stel ik een alternatief metametafysisch beeld voor dat niet ten prooi valt aan de problemen die ik voor het bestaande beeld signaleer.

De genoemde kritische zoektocht vindt grotendeels plaats in hoofdstuk 1 (hoewel er nog belangrijke toevoegingen aan gedaan worden in Deel II). Het metametafysische beeld dat ik zodoende boven tafel breng noem ik het *Moderne beeld*. In datzelfde hoofdstuk schets ik ook het alternatieve beeld, waarin mijn stelling dat begrippen essenties zijn een hoofdrol speelt. Dat beeld noem ik het *Aristotelische beeld*. Hoofdstukken 2 en 3 zijn vervolgens gewijd aan een preciezere verdediging van het Aristotelische beeld: in hoofdstuk 2 schets en verdedig ik de versie van essentialisme die daarbij hoort, en in hoofdstuk 3 schets ik de notie van conceptuele waarheid, die ook een belangrijke rol speelt. Aangezien het Aristotelische beeld de stelling omvat dat concepten essenties zijn, vormen hoofdstukken 2 en 3 als het ware twee zijden van één en dezelfde munt: essentiële waarheid is conceptuele waarheid. In deze samenvatting sta ik hoofdzakelijk stil bij de resultaten van hoofdstuk 1, omdat daarin de meest wezenlijke zaken geïntroduceerd worden. Mijn schets van hoofdstukken 2 en 3 is aanzienlijk korter.

Hoofdstuk 1: Het Aristotelische Beeld

In het leven van alledag, en voordat we ons uitgebreider gaan bezighouden met wetenschap en filosofie, nemen we de werkelijkheid simpelweg zoals die zich aan ons voordoet: we zijn *naïef realisten*, we beschouwen onszelf als wezens aan wie de werkelijkheid in de waarneming gegeven is. Zo gauw we echter beginnen deze naïeve positie kritisch te onderzoeken, gaan we eraan twijfelen. Ter illustratie noem ik hier enkele gangbare filosofische overwegingen in die richting. We merken bijvoorbeeld op dat onze zintuigen een hoogst belangrijke rol spelen voor onze visie op de werkelijkheid, we ontdekken dat die zintuigen gemakkelijk kunnen worden misleid, dat allerlei diersoorten gevoelig zijn voor aspecten van hun omgeving die door onze zintuigen niet kunnen worden geregistreerd (en andersom), enzovoorts. Dergelijke observaties suggereren dat wat we vanuit ons naïeve standpunt voor werkelijk hielden slechts het effect op ons organisme is van de ware werkelijkheid, die dus achter de sluier van zintuiglijke indrukken schuilgaat. Aan de andere kant merken we ook op dat de begrippen waarmee we de werkelijkheid benaderen een doorslaggevende rol spelen in onze beleving ervan: mijn vriend, die een ervaren ecoloog is, neemt over het algemeen veel meer waar dan ik als we eens over de hei wandelen, simpelweg omdat ik niet over het begripsmatige arsenaal beschik dat hij zich van beroepswege eigen heeft gemaakt. En als we wat verder van huis kijken, ontdekken we dat er grote verschillen zijn tussen de begrippenkaders van verschillende culturen en volkeren: we ontdekken hoe intiem de verbinding tussen ons begrippenkader en onze visie op de werkelijkheid is, zowel in ons begrip als in onze waarneming ervan. Ook hier lijkt de les te zijn dat wat we vanuit het naïeve standpunt voor werkelijk hielden niet de echte werkelijkheid is: het wordt mede gevormd door ons eigen begrippenkader. Hoe het resultaat van die vorming zich dan verhoudt tot de werkelijkheid *op zichzelf* wordt zo een prangende vraag.

Aldus komen we uit op de voorstelling dat er twee niveaus van werkelijkheid zijn: er is de werkelijkheid zoals die zich aan ons voordoet, en er is de werkelijkheid zoals die op zichzelf is. De werkelijkheid-op-zich ligt wel op de een of andere wijze ten grondslag aan de werkelijkheid-voor-ons, maar die laatste is gekleurd door de eigenaardigheden van ons waarnemingsvermogen en van ons denkvermogen. Deze fundamentele scheiding tussen werkelijkheid-op-zich en werkelijkheid-voor-ons vormt de kern van wat ik het *Moderne beeld* noem.[393] Bezien tegen de achtergrond van dit metametafysische beeld is het de opgave van de metafysica om tot een

[393] Ik noem het 'Modern' omdat het zich voornamelijk sinds de vroeg-moderne periode in de wijsbegeerte, zo vanaf de vijftiende eeuw, heeft doen gelden.

ware theorie van de werkelijkheid-op-zich te komen, en de relatie daarvan met de werkelijkheid-voor-ons vast te stellen. Omdat wat we in eerste instantie voor ons hebben de werkelijkheid-voor-ons is, is de enige mogelijkheid die we hebben de werkelijkheid-op-zich te *reconstrueren* aan de hand van de werkelijkheid-voor-ons. Het doel van de metafysica is dus, beschouwd op basis van dit Moderne beeld, om de werkelijkheid-voor-ons te *verklaren* door middel van een theorie over hoe de werkelijkheid-op-zich fundamenteel in elkaar steekt.

Het Moderne beeld biedt ruimte voor drie grondverschillende opvattingen over de mogelijkheden en opgaven van de metafysica. De eerste noem ik *metafysisch realisme*, de tweede *metafysisch anti-realisme*, en de derde *metafysisch skepticisme*. Dit komt ruwweg overeen met wat je in grote delen van de hedendaagse metafysische debatten terugvindt: daar is inderdaad een fundamentele tegenstelling te ontwaren tussen metafysisch realisten en metafysisch anti-realisten—skeptici zijn minder duidelijk vertegenwoordigd.[394] De wijze waarop ik hierboven de taak van de metafysica volgens het Moderne beeld schetste—namelijk als de opgave om tot een theorie van de werkelijkheid-op-zich te komen die dan de werkelijkheid-voor-ons moet verklaren—ligt dichtbij metafysisch realisme: realisten menen dat een dergelijk soort metafysica mogelijk is. De metafysisch skepticus is het met de realist eens wat betreft de doelstelling van de metafysica, maar beargumenteert dat het onmogelijk is om die doelstelling te behalen, in wezen om de volgende reden: elke theorie die wij verzinnen is toch uiteindelijk product van ons denken, wordt ingepast in ons begrippenkader, en blijft daarmee altijd gevangen in de werkelijkheid-voor-ons. De kloof, zo meent de skepticus, is niet te overbruggen. De anti-realisten, tenslotte, zijn het weliswaar met de skepticus eens dat die kloof niet te overbruggen is, maar vinden dat dat laat zien dat we ons de doelstelling van de metafysica verkeerd voorstellen. In plaats van het overbruggen van de kloof, zo zeggen anti-realisten, moet de mefafysica zich beperken tot het opstellen van een visie op de werkelijkheid-voor-ons.[395] De metafysisch realist zal een dergelijke afzwakking van het metafysische project natuurlijk krachtig van de hand wijzen: we willen immers weten hoe de werkelijkheid-op-zich in elkaar steekt, en niet slechts hoe de werkelijkheid-voor-ons

[394] Ik onderzoek metafysisch realisme in §1.1.1 en metafysisch anti-realisme in §1.1.2. In §1.4 beschouw ik de pessimistische positie van Barry Stroud, die dicht in de buurt komt van het skepticisme zoals hier bedoeld.

[395] Er zijn veel soorten anti-realisme. Eén optie, die we Kantiaans kunnen noemen, bestaat erin te denken dat onze menselijke geest een noodzakelijke ordening aan onze ervaring oplegt, zodat het die ordening is waar de metafysica om draait. Een andere optie, die we conventionalisme kunnen noemen, bestaat erin te denken dat we uiteindelijk op puur pragmatische gronden een of ander begrippenkader moeten kiezen, wat dan de bril wordt waardoorheen we de werkelijkheid zien. De metafysica draait dan om het uiteenzetten van de eigenschappen van die willekeurig gekozen bril.

in elkaar steekt.

Het volgende overzicht laat de verhoudingen tussen de genoemde drie op het Moderne beeld gebaseerde visies op metafysica zien:

	Skepticisme	*Anti-Realisme*	*Realisme*
Kunnen we tot inzicht komen in de werkelijkheid-op-zich?	Nee	Nee	Ja
Is het doel van de metafysica zulk inzicht in de werkelijkheid-op-zich?	Ja	Nee	Ja

De skepticus kunnen we het beste als de uitdager beschouwen, waarop de realist en de anti-realist verschillend antwoorden. Maar beide antwoorden zijn problematisch. Omdat het Moderne beeld volledig open laat hoe de werkelijkheid-op-zich eruit ziet, staat het de metafysisch realist vrij om de wildste theorieën daarover te bedenken en met behulp van slimme analyses te laten zien hoe we de werkelijkheid-voor-ons in die termen kunnen begrijpen.[396] Dat maakt het meteen vrijwel onmogelijk om tot een gefundeerd oordeel te komen over welke van dergelijke (wilde en minder wilde) theorieën de juiste is. In wezen is de realist bezig met een *abductieve* vorm van verklaring: hij zoekt naar de theorie van de werkelijkheid-op-zich die de meest waarschijnlijke verklaring biedt voor hoe de werkelijkheid-voor-ons eruit ziet—alleen biedt de manier waarop het vraagstuk bekeken wordt te weinig houvast om tot een oordeel te komen over de vraag of de ene verklaring nu waarschijnlijker is dan de andere of niet.

Ook het anti-realisme is uiteindelijk niet bevredigend. We duidden het probleem al aan: de anti-realist wil de doelstelling van de metafysica veranderen, maar heeft voor die verandering wel nodig dat we vanuit het Moderne beeld blijven denken. Waar de metafysica zich op moet richten is immers, volgens de anti-realist, de werkelijkheid-voor-ons *in plaats van* de werkelijkheid-op-zich.

Daarom is het verstandig om te bekijken of er geen alternatief te vinden is voor het Moderne beeld, dat ten grondslag ligt aan het geschetste drieluik van visies op de metafysica. Want het is natuurlijk heel goed mogelijk dat de aangeduide problemen uiteindelijk op dat Moderne beeld terug te voeren zijn. Het gezochte alternatief is redelijk eenvoudig te construeren, namelijk door de scheiding tussen de werkelijkheid-voor-ons en de werkelijkheid-op-zich, waar het Moderne beeld om

[396] Een sprekend voorbeeld is het 'blobjectivisme' van Horgan en Potrč [2008], de visie dat er maar één object bestaat (het blobject) en dat alle waarheden die schijnbaar over afzonderlijke objecten gaan (zoals tafels, bomen, sterren, en mensen) te analyseren zijn als waarheden over dat ene blobject.

draait, van de hand te wijzen en dat consequent door te denken. Zodoende kom ik tot het *Aristotelische beeld*.

In tegenstelling tot het Moderne beeld, bestaat er in het Aristotelische beeld dus geen fundamentele scheiding tussen de werkelijkheid-voor-ons en de werkelijkheid-op-zich. Er is één werkelijkheid, en die doet zich aan ons voor in waarneming én begrip. De begrippen waarmee wij denken constitueren zelf de werkelijkheid: het zijn de essenties van wat er in de werkelijkheid zoal bestaat. Dus, als je de juiste begrippen vat wanneer je je met één of ander stukje werkelijkheid geconfronteerd ziet dan heb je daadwerkelijkheid het inzicht bereikt waar het daar om gaat. Daarmee doet het Aristotelische beeld recht aan de intuitie van de metafysisch realist: we kunnen inderdaad tot inzicht in de werkelijkheid komen. Maar, in tegenstelling tot wat de realist denkt, is er geen *verdere* stap nodig die de begrippen die we in zo'n situatie vatten verbindt met een begrips-loze, verborgen werkelijkheid-op-zich. En daarmee doet het Aristotelische beeld tevens recht aan de intuitie van de metafysisch anti-realist: het is onmogelijk om 'uit' ons begrippenkader te stappen en de begrips-loze 'echte' werkelijkheid te aanschouwen. De reden daarvoor is echter niet, zoals de anti-realist beweert, dat we gevangen zouden zijn in een begrijpelijke werkelijkheid-voor-ons die gescheiden is van de on-begrijpelijke werkelijkheid-op-zich—de reden is eenvoudigweg dat we met onze begrippen al middenin de (enige) werkelijkheid staan. De hele notie van een begrips-loze werkelijkheid is een misvatting.

Wat overblijft is het interessante maar metafysisch beschouwd onschuldige gegeven dat we feilbare wezens zijn: we kunnen ons vergissen, de verkeerde begrippen vatten, of de juiste begrippen onvolledig vatten. Want omdat onze begrippen niet het *product* van ons denken zijn, maar er juist door *gevat* worden, kunnen we inzien dat er een discrepantie kan bestaan tussen het begrip zelf en wat ik daarvan weet te vatten. Die discrepantie *kan* bestaan, maar *hoeft* niet te bestaan—in succesvolle gevallen is hij afwezig. We moeten de begrippen dus niet verwarren met datgene wat ik ervan weet te vatten op een zeker moment, maar daarmee zijn die twee nog niet fundamenteel verschillend.

In het Aristotelische beeld gaapt er geen kloof tussen de werkelijkheid-voor-ons en de werkelijkheid-op-zich. Er gaapt geen kloof tussen wat begripsmatig is en wat werkelijk is. Het kan daarom niet de taak van de metafysica zijn die kloof te overbruggen. In plaats daarvan is het de taak van de metafysica om de begrippen waarmee we tot inzicht in de werkelijkheid komen zogezegd tot op de bodem te doordenken, onderlinge samenhangen bloot te leggen, eventuele fouten en onzuiverheden te ontmaskeren, en zodoende tot een overkoepelend inzicht te komen in de

samenhang der dingen. Vanuit dit gezichtspunt bekijk ik in deel II een aantal cruciale metafysische vraagstukken die in het hedendaagse debat leven—over natuurwetten, causaliteit, tijd, en leven—om tot een bescheiden, eerste schets van de genoemde samenhang der dingen te komen.

Maar eerst nog het volgende. Het belangrijkste obstakel voor de aanvaardbaarheid van het Aristotelische beeld is de opmerkelijke rol die het aan begrippen toekent. Hoewel dat naar mijn indruk alleen een obstakel is zolang je vasthoudt—impliciet of expliciet—aan het Moderne beeld, is het wel zaak om te laten zien hoe begrippen nu precies *zowel* denkend te vatten *alsook* vormend onderdeel van de werkelijkheid kunnen zijn. Dat is de taak waar ik hoofdstukken 2 en 3 aan gewijd heb.

Hoofdstuk 2: Essentiële Waarheid

Essentialisme maakt deel uit van het Aristotelische beeld—het is het deel dat duidelijk maakt hoe begrippen de werkelijkheid kunnen vormen. Een belangrijke rol is in dit verband weggelegd voor die begrippen die ik *ontologisch fundamenteel* noem.[397] Zulke begrippen vormen het uiteindelijke antwoord op de Aristotelische vraag *wat is dit?*, toegepast op wat dan ook. Ze zijn als volgt te onderscheiden van niet-fundamentele begrippen: de niet-fundamentele begrippen zijn reduceerbaar tot zulke fundamentele begrippen. Neem bijvoorbeeld het begrip *vrijgezel*: een vrijgezel *is* eenvoudigweg een mens die volwassen is en geen relatie heeft. Het fundamentele begrip hier is *mens*, het begrip *vrijgezel* is daarvan afgeleid. Een niet-fundamenteel begrip is dus terug te brengen tot een fundamenteel begrip plus bepaalde verdere condities. De fundamentele begrippen zijn die begrippen die *niet* op die manier te reduceren zijn. Het punt is nu dat alles wat volgt uit het feit dat een gegeven ding onder een bepaald fundamenteel begrip valt, essentieel is voor dat ding. Het fundamentele begrip in kwestie *is* de essentie van dat ding.

De belangrijkste conclusies van dit hoofdstuk zijn niet alleen van toepassing op metafysica tegen de achtergrond van het Aristotelische beeld. Want ook de metafysisch realist die van het Moderne beeld uit gaat wil tot een zekere lijst van ontologisch fundamentele begrippen komen. Voor hem volgt er echter nog een stap: die van de begrippen naar de dingen zelf. Hij stelt dat de begripsmatige samenhang die een aanhanger van het Aristotelische beeld als zodanig accepteert slechts een afspiegeling is van een gehypothetiseerde metafysische samenhang in de werkelijkheid-op-zich. Die complicatie is, bezien vanuit het Aristotelische beeld,

[397] 'Ontologisch' wil hier zoveel zeggen als: 'wat het zijn betreft'.

volledig overbodig. De anti-realist is het daarmee eens: hij zegt ook dat de hele idee van een dergelijke metafysische samenhang onzinnig is (omdat het immers toch weer meer theorie, en dus begripsmatig is)—maar stelt daarentegen weer dat die begripsmatige samenhang niets met de werkelijkheid-op-zich van doen heeft (en dat we er desondanks tevreden mee moeten zijn). Ook deze complicatie is, bezien vanuit het Aristotelische beeld, volledig overbodig.

Hoofdstuk 3: Conceptuele Waarheid

In mijn uiteenzetting over essentialisme maakte ik gebruik van het begrip *gevolg*, van het idee dat bepaalde essentiële conclusies uit een gegeven fundamenteel begrip kunnen *volgen*. Het is van belang om dat idee beter te begrijpen: daarmee komen we weer een stapje dichter bij wat ik bedoel met de frase 'begrippen zijn essenties'. Waar ik in hoofdstuk 2 dus de nadruk legde op essenties, ligt in dit hoofdstuk de nadruk op begrippen.

De studie van gevolgtrekking, consequentie, of geldig redeneren vormt al sinds de introductie ervan door Aristoteles de focus van de discipline die we *logica* noemen. De cruciale notie is die van *geldigheid*: een redenering, bestaande uit een aantal premissen en een conclusie, is geldig dan en slechts dan als de waarheid van de premissen de waarheid van de conclusie *garandeert*. Bij een geldige redenering is het *onmogelijk* dat alle premissen waar zijn terwijl de conclusie onwaar is; de conclusie *moet* waar zijn als de premissen allemaal waar zijn. Dat maakt geldigheid tot een modale notie (die dus niet slechts iets zegt over wat nu eenmaal het geval is, maar ook over wat noodzakelijk en (on)mogelijk is).

De werkwijze van logici is, eveneens al sinds Aristoteles, om de wetten van geldigheid te achterhalen door te abstraheren. We constateren bijvoorbeeld dat uit de premissen 'Alle mensen zijn sterfelijk' en 'Socrates is een mens' de conclusie 'Socrates is sterfelijk' volgt, en we merken meteen op dat dat geldt voor alle redeneringen van de vorm 'Alle A's zijn B; a is A; dus a is B'. Laten we zo'n vorm een *formalisering* noemen van de oorspronkelijke redenering. Natuurlijk kun je op verschillende manieren tot een dergelijke formalisering komen—een andere zou bijvoorbeeld zijn 'Alle mensen zijn A; Socrates is een mens; dus Socrates is A'.

Het ligt nu voor de hand te denken dat één bepaalde keuze van formalisering *de juiste* is, namelijk die, die de *logische vorm* van de redenering in kwestie isoleert. Vrijwel iedereen zal inzien dat de eerste van onze voorbeeld-formaliseringen in dat opzicht beter scoort dan de tweede: de geldigheid van de voorbeeld-redenering

hangt immers niet af van het gegeven dat we het over mensen en over Socrates hebben, we kunnen dus net zo goed van die inhoud afzien. In de analytische traditie waaraan mijn onderzoek zich oriënteert, die sterk empiristische wortels heeft, heerst de gedachte dat er een fundamentele scheiding is tussen de *vorm* van redeneringen (het begripsmatig raamwerk) en hun *inhoud* (die empirisch is). De vorm is, volgens deze gedachte, datgene wat overblijft als je *alle* inhoud weglaat. Ik beargumenteer daarentegen dat dit een misvatting is. Wat we doen als we een formalisering opstellen zoals we hierboven deden, is het volgende. We laten *delen* van de inhoud van een redenering weg, vervangen die door schematische letters en dergelijken, zodat we alleen dat deel van de oorspronkelijke inhoud overhouden dat bepalend is voor de geldigheid van de oorspronkelijke redenering. Het kan heel verschillend zijn welke aspecten van de inhoud daartoe geïsoleerd moeten worden, voor verschillende redeneringen. Neem bijvoorbeeld de redenering die uit de premisse 'Jan is een vrijgezel' de conclusie 'Jan is ongetrouwd' trekt: formaliseren we die als ''*a* is *A*, dus *a* is *B*' dan zien we dat die vorm ook ongeldige instanties heeft (bijvoorbeeld: 'Jan is een mens, dus Jan is een paard'). Het verband tussen het begrip *vrijgezel* en het begrip *ongetrouwd* (of, misschien beter, tussen *vrijgezel*, *getrouwd* en *ontkenning*) speelt een rol voor de geldigheid van de genoemde redenering.

Mijn stelling in dezen is dat de fundamentele notie voor de logica de notie van *waarheid gegarandeerd door inhoud* is. De logica bestudeert dus hoe inhoud waarheid garandeert. Dat de logische discipline zich daarbij tot bepaalde *soorten* van inhoud beperkt, namelijk inhoud van heel algemene aard zoals de gebruikelijke voegwoorden ('en', 'of', 'als ... dan', etc.) en kwantoren ('sommige', 'alle')—maar overigens ook temporele, modale, en andere noties bestudeert—is heel begrijpelijk (en vruchtbaar), en daarom kan uit pragmatisch oogpunt voor zulk onderzoek de term 'logica' en 'logisch' gereserveerd worden. De logische vorm van een uitspraak bestaat dan uit precies die aspecten van de inhoud ervan die de logica zich tot onderwerp maakt. In dat geval kan er echter geen sprake van zijn dat de logica zeggenschap heeft over de volledige reikwijdte van het begrip geldigheid: het beperkt zich immers tot geldigheid op basis van een afgebakend stuk mogelijke inhoud, terwijl geldigheid een notie is die op *alle* soorten inhoud kan stoelen. Wat logisch geldig is, in deze afgebakende zin van 'logisch', is dus een beperking op wat geldig is in de ruimere zin. Ik noem die ruimere notie van geldigheid *conceptuele geldigheid*.

Een broertje van logische geldigheid is *logische waarheid*. Een uitspraak of gedachte is logisch waar dan en slechts dan als zijn inhoud gegarandeerd wordt door zijn logische vorm—dat wil zeggen, door dat deel van de inhoud waar de logica als

discipline zich mee bezig houdt. Een voorbeeld is 'Het regent of het regent niet': dat is waar puur vanwege de wijze waarop twee keer dezelfde uitspraak gecombineerd wordt met 'of' en 'niet'. Mijn eerdere gedachtengang volgend komen we gemakkelijk tot een ruimere notie, analoog aan conceptuele geldigheid, die ik *conceptuele waarheid* noem: een uitspraak of gedachte is conceptueel waar dan en slechts dan als zijn waarheid gegarandeerd wordt door zijn inhoud (nu zonder beperking). Het is deze notie van conceptuele waarheid waar de versie van essentialisme die ik in hoofdstuk 2 verdedigd heb op stoelt: essentiële waarheden over het een of andere ding zijn die waarheden die volgen uit het fundamentele begrip waar dat ding onder valt; de waarheid daarvan wordt gegarandeerd door het fundamentele begrip in kwestie—het zijn dus conceptuele waarheden.

Het geschetste idee van conceptuele waarheid en geldigheid, van waarheid en geldigheid gegarandeerd door inhoud, voert tot de gedachte dat er meer af te bakenen soorten van inhoud kunnen zijn dan die we normaal gesproken in de logica vinden—in het bijzonder soorten van inhoud, clusters van begrippen, die vanuit een metafysisch oogpunt interessant zijn. Dat is de gedachte waar Deel II van dit proefschrift mee begint.

Deel II: Metafysica

De metafysische onderzoekingen die ik in dit tweede deel onderneem, gebaseerd op het Aristotelische beeld zoals geschetst en ontwikkeld in Deel I (hierboven samengevat), gaan uit van een observatie die Michael Thompson in zijn [2008] maakt. Die observatie is dat er verschillende vormen van *predicatie* bestaan, dat wil zeggen, verschillende wijzen waarop je iets van iets kunt zeggen. Om zulke vormen van predicatie formeren zich clusters van verwante begrippen: elke vorm van predicatie kent een bijpassende klasse van eigenschappen en relaties en een bijpassend soort individuele objecten, en kenmerkt zich door typische vormen van conceptuele geldigheid en conceptuele waarheid. Bovendien laten de conceptuele clusters die je zodoende kunt ontwaren zich hiërarchisch ordenen—daarom spreek ik ook van *conceptuele niveaus*. Thompson illustreert deze gedachte door middel van een beeld: dat van een conceptuele versnellingsbak met diverse conceptuele versnellingen ('gears'): bij het schakelen van de ene naar de andere versnelling blijven bepaalde dingen gelijk, maar andere dingen worden anders—de begripsmatige tandwielen grijpen anders in elkaar, om bij het beeld te blijven. In het bijzonder blijken de hogere conceptuele niveaus meer differentiatie en complexiteit te vertonen dan de lagere niveaus.

In dit proefschrift bespreek en verdedig ik drie conceptuele niveaus, zonder overigens daarmee te willen zeggen dat er niet nog meer te onderscheiden zijn. Hoofdstuk 4 biedt een overzicht van deze drie conceptuele niveaus, zonder daar al meteen de uiteenzetting met relevante hedendaagse metafysische debatten bij te betrekken. Hoofdstukken 5, 6, en 7 zijn juist aan een dergelijke uiteenzetting gewijd, waarbij hoofdstuk 5 zich met name op het eerste conceptuele niveau richt, hoofdstuk 6 op het tweede conceptuele niveau, en hoofdstuk 7 op het derde conceptuele niveau. Door zo de relevante conceptuele niveaus in hun samenhang met cruciale vraagstukken uit de hedendaagse metafysica uit te werken, laat ik zien dat de relevante conceptuele niveaus een bepaalde zelfstandigheid van elkaar, een zekere onafhankelijkheid genieten—met name in de zin dat hogere conceptuele niveaus onreduceerbaar zijn tot lagere niveaus. Voor deze samenvatting beperk ik mij voornamelijk tot de inhoud van hoofdstuk 4, en geef slechts een kleine proeve van wat in de verdere hoofdstukken besproken wordt.

Hoofdstuk 4: Conceptuele Niveaus

Laten we beginnen met een schematisch overzicht van de drie conceptuele niveaus waarop ik me in dit tweede deel concentreer:

	Eerste niveau	**Tweede niveau**	**Derde niveau**
vorm van predicatie:	tempusloos	tempus/aspect	normatief
van toepassing op:	entiteit	object	organisme
vallen onder:	type	natuurlijke soort	levensvorm
predicaten:	eigenschap	toestand/proces	levensproces
modus van zijn:	bestaan	voortduren	leven
activiteit:	—	processen	levensprocessen
karakteristiek:	wiskundig	causaal/temporeel	teleologisch
	Humeanisme	*anti-Humeanisme*	

Het eerste conceptuele niveau is het niveau van de meest eenvoudige vorm van predicatie. Dit conceptuele niveau vindt bij uitstek toepassing in gedachten van wiskundige aard. Hier vinden we de pure categorie van een *entiteit*, in de abstracte zin van *iets waar je iets van kunt prediceren*—bijvoorbeeld een zekere driehoek. En we vinden de pure categorie van een eigenschap of relatie, eveneens in de abstracte zin van *iets wat je van een entiteit of entiteiten kunt prediceren*—bijvoorbeeld de eigenschap *is rechthoekig*, of de relatie *is congruent met*. De vorm van predicatie waar het hier om gaat noem ik *eigenschapspredicatie*. Eigenschapspredicatie is ongedifferentieerd: het verbindt eenvoudigweg een eigenschap met een entiteit, of een relatie met meerdere

entiteiten.[398] Ter illustratie: zeggen we van een bepaalde driehoek op het Euclidische vlak dat die rechthoekig is, dan is er geen sprake van dat die driehoek ooit *niet* rechthoekig was of zal zijn (of had kunnen zijn): eigenschapspredicatie van het eerste conceptuele niveau is *atemporeel*, kent geen differentiatie in tempora (en is tevens *amodaal*, kent geen onderscheid tussen actualiteit en mogelijkheid).

Zoals ik beargumenteerde in hoofdstuk 2 valt alles wat bestaat uiteindelijk onder het één of andere fundamentele begrip, en dat is ook hier zo. Ik noem de soorten fundamentele begrippen van het eerste concepuele niveau *types*. Zeggen dat een entiteit onder een bepaald type valt is iets anders dan zeggen dat het de een of andere eigenschap heeft; deze speciale vorm van predicatie noem ik *soortpredicatie*. Soortpredicatie is eveneens een ongedifferentieerde vorm van predicatie, en komt bovendien op alle conceptuele niveaus voor. Op het eerste conceptuele niveau verbindt het entiteiten met types (of types met types). Bijvoorbeeld: 'deze figuur is een driehoek' (of: 'de driehoek is een polygoon'). Ware uitspraken met soortpredicatie zijn altijd conceptueel waar.

Alle elementen die we nu in hun voorkomen op het eerste conceptuele niveau geschetst hebben, komen ook voor op het tweede conceptuele niveau, maar dan getransformeerd en gedifferentieerd. In plaats van entiteiten vinden we hier de pure categorie van *fysieke objecten*, die bestaan doordat ze door de tijd voortduren ('persisteren'): bijvoorbeeld een bepaalde bal. De fundamentele begrippen waaronder zulke fysieke objecten vallen noem ik *natuurlijke soorten*; bijvoorbeeld *bal*, of *ster*, of *goud*. In plaats van eigenschappen hebben deze fysieke objecten *toestanden* en *activiteiten*: hier zien we al enige differentiatie. De vorm van predicatie die deze zaken met elkaar verbindt kent een tempus (in het geval van toestanden) en daarnaast ook een aspect (in het geval van activiteiten). Bijvoorbeeld: 'de bal is rood', 'de bal was rood' (tempus), en: 'de bal is aan het rollen' (progressief aspect), 'de bal heeft gerold' (perfectief aspect). Tijd en oorzakelijkheid zijn de categorieën die bij deze kenmerken van predicatie op het tweede conceptuele niveau horen.

Dat fysieke objecten activiteiten kunnen ontplooien is iets dat op zichzelf verdere onderscheidingen vereist. Want fysiek objecten kunnen verschillende activiteiten ontplooien—ze hebben *causale vermogens* daartoe. Een causaal vermogen is het vermogen om een bepaalde activiteit te ontplooien wanneer de juiste omstandigheden zich voordoen. Er bestaat op het tweede conceptuele niveau een onderscheid tussen wat slechts mogelijk is en wat verwerkelijkt is (of wordt). De activiteiten zelf zijn op

[398] Eigenschapspredicatie is de vorm van predicatie waar predicatenlogica mee werkt: het is de vorm van predicatie die zich, in Fregeaanse termen, als functie-applicatie laat beschrijven.

hun beurt te beschouwen als *processen*, die zelf ook weer onder natuurlijke soorten vallen en bepaalde toestanden kennen. Al deze onderscheidingen en differentiaties komen op het eerste conceptuele niveau niet voor.

Op het derde conceptuele niveau vinden we een verdere transformatie van dit al veel rijkere conceptuele cluster. De pure categorie van een *organisme* neemt hier de plaats in van fysieke objecten: het bestaan van een organisme ligt niet slechts in zijn voort*duren* in de tijd maar vereist een voort*leven* in de tijd. De fundamentele begrippen waar zulke organismes onder vallen noem ik *levensvormen* (*life-forms*). Ook de categorie van processen maakt plaats voor iets nieuws, wat ik *levensprocessen* noem. De levensprocessen waar een organisme in verwikkeld is spelen een heel bepaalde, zinvolle rol voor het leven van dat organisme en worden dan ook gekenmerkt door een bepaalde vorm van normativiteit: organismes *behoren* de levensprocessen uit te voeren die in hun levensvorm vervat liggen. Die levensprocessen hangen dan ook op een harmonieuze wijze samen. Levensvormen omvatten zodoende een zeer dynamisch principe van organisatie dat zichzelf uitdrukt in de wijze waarop hun instanties (individuele organismes) de hun beschikbare materie organiseren, zich ontwikkelen, groeien, zich tegen schadelijke invloeden beschermen, en zich voortplanten (onder andere). Deze normativiteit kenmerkt ook de vorm van predicatie waarmee toestanden, processen en levensprocessen aan organismes toegeschreven kunnen worden. Het typische van deze normativiteit is dat het kan voorkomen dat een organisme *niet* aan de kenmerkende eisen van zijn levensvorm voldoet—zo kan bijvoorbeeld een paardenkastanje, die in de lente zijn typische handvormige bladeren *hoort* te ontvouwen, daarin toch falen doordat hij geplaagd wordt door de paardenkastanjemineermot, of doordat hij lijdt aan de kastanjebloedingsziekte, of doordat hij kunstmatig in wintertoestand gehouden wordt (of natuurlijk doordat hij eenvoudigweg omgezaagd wordt).

We begrijpen grote delen van de werkelijkheid met behulp van alle drie deze conceptuele clusters—en het ligt voor de hand te denken dat er nog meer conceptuele niveaus zijn behalve die waar ik me in dit proefschrift toe beperk. Vanuit het Aristotelisch uitgangspunt gedacht bestaat metafysica erin dat we de algemene conceptuele structuren die we voor een goed begrip van de werkelijkheid nodig hebben op zichzelf beschouwen en uitdiepen. In tegenstelling daartoe begrijpen metafysici die van het Moderne beeld uitgaan hun opgave heel anders. Voor mijn onderzoekingen in dit tweede deel zijn vooral de metafysisch realisten van belang: zij willen tot een theorie over de werkelijkheid-op-zich komen waarvan ze dan kunnen aantonen dat die aan de werkelijkheid-voor-ons ten grondslag ligt.

Met het onderscheid tussen de diverse conceptuele niveaus bij de hand kunnen we zulke metafysisch realisten grofweg indelen in diverse kampen. Er zijn metafysici die menen dat alleen het soort conceptuele structuur dat we op het eerste niveau vinden overeenkomt met de werkelijkheid-op-zich. Zij verdedigen een of andere versie van wat ik *eerste-niveau metafysica* noem: in wezen draait het om een reductie van alle conceptuele rijkdom die we op hogere niveaus vinden tot een simulatie daarvan in termen van het eerste conceptuele niveau. Filosofen die zich bij een Humeaanse zienswijze thuisvoelen neigen meestal tot een eerste-niveau metafysica.[399]

Er zijn daarnaast ook metafysici die zich niet willen beperken tot het eerste conceptuele niveau: zij menen dat de werkelijkheid-op-zich rijker is dan dat. Zoals te verwachten valt keren zij zich tegen de Humeanen en zijn dan ook bekend onder de naam *anti-Humeanen*. De meesten van deze metafysici willen zich echter wél beperken tot het tweede niveau, en gaan er dus vanuit dat de eigenaardigheden van het derde conceptuele niveau op een of andere wijze terug te voeren zijn, te reconstrueren zijn, met behulp van puur tweede-niveau denkwerk—met andere woorden, dat het derde conceptuele niveau slechts bij de werkelijkheid-voor-ons hoort en niet bij de werkelijkheid-op-zich. Dat is wat ik *tweede-niveau metafysica* noem.

In hoofdstuk 5 bespreek ik eerste-niveau metafysica en de samenhang daarvan met Humeaans gedachtengoed. Een typisch voorbeeld van een eerste-niveau denker is David Lewis; ik sta dan ook vrij uitvoerig bij zijn metafysische programma stil. En bij wijze van voorbeeld ga ik dieper in op de discussie over natuurwetten—Humeaanse visies op natuurwetten genieten al sinds het begin van de analytische wijsbegeerte een aanzienlijke populariteit.

In hoofdstuk 6 bespreek ik de sterke oppositie tegen eerste-niveau metafysica, en bespreek daarbij uitvoeriger de hedendaagse discussie over causaliteit en over tijd—de twee begrippen die een zo centrale rol spelen in het tweede conceptuele niveau. Wat ik zodoende laat zien is dat het niet volstaat om slechts een gedeelte of aspect van het tweede conceptuele cluster te willen inbouwen in een metafysische theorie die zich overigens tot het eerste niveau beperkt: het resultaat is dan steeds toch een reductieve theorie, hoewel het reductieve karakter ervan vaak verborgen is onder een laag anti-reductieve retoriek. Een voorbeeld van zo'n onsuccesvolle anti-Humeaanse visie is die van David Armstrong op oorzakelijkheid: omdat de rest van zijn metafysische plaatje duidelijk binnen de kaders van het eerste conceptuele niveau blijft, lukt het hem niet om daadwerkelijk tot een tweede-niveau begrip van

[399] Maar Humeanisme en eerste-niveau metafysica vallen niet exact samen, zoals ik in §5.2.4 beargumenteer.

oorzakelijkheid te komen—ondanks zijn pogingen om zich tegen Humeaanse visies op oorzakelijkheid af te zetten. Daarmee wordt ook duidelijker dat het tweede conceptuele niveau inderdaad als een samenhangend geheel op zichzelf te beschouwen is, onafhankelijk van het eerste conceptuele niveau.

In hoofdstuk 7 tenslotte gaat het om het derde conceptuele niveau. Waar ik me in hoofdstukken 5 en 6 uiteen kon zetten met de bestaande (en beruchte) discussie tussen relatief duidelijke opponenten—Humeanen en anti-Humeanen—ligt de zaak hier anders: wat de levende natuur betreft overheerst de gedachte dat die uiteindelijk terug te voeren moet zijn op causale processen die ook onafhankelijk van hun voorkomen in organismen te begrijpen zijn. Wat dat precies betekent is weliswaar onderwerp van een levendig debat over reductionisme in de filosofie van de biologie, maar ook diegenen die zich in dat debat tegen reductie afzetten omarmen veelal toch de genoemde gedachte in een of andere vorm, of komen niet verder dan enkele onduidelijke opmerkingen over veroorzaking 'van bovenaf' en emergentie (de these dat gehelen eigenschappen kunnen hebben die niet verklaarbaar zijn uit het samenspel van de delen). Mijn stelling is dat de genoemde gedachte een uitdrukking is van tweede-niveau metafysica, en dus voortkomt uit het Moderne beeld: de veronderstelling is dat het derde conceptuele niveau slechts van toepassing is op wat zich als levende natuur aan ons voordoet (een stukje werkelijkheid-voor-ons dus), terwijl de werkelijkheid-op-zich die daaraan ten grondslag ligt puur vanuit het tweede conceptuele niveau te begrijpen moet zijn—dus zonder de normativiteit en teleologie die zo typisch zijn voor het derde conceptuele niveau. Deze visie is dus gecommitteerd aan een reductieve visie op de levende natuur. Bij wijze van voorbeeld bespreek ik uitvoeriger een nauwkeurig uitgewerkte versie van zo'n reductieve visie, namelijk de definitie van *functie* (in de biologische zin) die Ruth Millikan ontwikkeld heeft. Ik beargumenteer dat die definitie er niet in slaagt om datgene te vangen wat zich in termen van het derde conceptuele niveau laat uitdrukken. Ook hier is de conclusie dat het derde conceptuele niveau inderdaad als een samenhangend geheel op zichzelf te beschouwen is, los van het eerste conceptuele niveau. Ik sluit hiermee aan bij de visie die Michael Thompson op leven ontwikkelt in deel I van zijn [2008].

Hoofdstuk 5: Het Eerste Conceptuele Niveau

Het eerste conceptuele niveau onderscheidt zich aanzienlijk van het tweede conceptuele niveau, met name ook wat een aantal metafysisch belangrijke noties betreft. Waar op het tweede conceptuele niveau begrippen als causaliteit en tijd een be-

langrijke rol spelen, zijn die volledig afwezig op het eerste conceptuele niveau. Het resultaat is dat metafysici die zich tot het eerste conceptuele niveau willen beperken—eerste-niveau metafysici dus—voor al dat soort begrippen een reductieve theorie moeten opstellen. Waar het tweede conceptuele niveau een bepaald stuk metafysica al *omvat*, moet dat op het eerste conceptuele niveau *gesimuleerd* worden.

Een goed voorbeeld van zulke simulaties is te vinden in het werk van David Lewis. Zijn metafysische programma bestaat uit een reeks vernuftig in elkaar passende simulaties van hogere-niveau fenomenen, resulterend in een indrukwekkende overkoepelende metafysische theorie. Ik geef hier slechts twee van de interessantere ingrediënten weer—namelijk die waar ik ook in mijn samenvatting van hoofdstuk 6 op terugkom: tijd en de aard van fysieke objecten.

Een eerste-niveau simulatie van *tijd* bestaat in het typische geval (zoals bij Lewis) uit drie elementen: *eternalisme*, *B-theorie*, en *perdurantisme* (de tegenhangers *presentisme*, *A-theorie* en *endurantisme* bespreek ik in hoofdstuk 6). Onder eternalisme versta ik de gedachte dat tijd, net als ruimte, een reeks van locaties biedt waar de dingen en de gebeurtenissen zich kunnen bevinden. De notie van een *temporele locatie* is een typisch eternalistische notie. B-theorie is de gedachte dat een in de tijd gesitueerde gedachte (zoals bijvoorbeeld 'ik schrijf een samenvatting') bestaat uit een *tijdloze* gedachte die op een bepaalde temporele locatie gesitueerd wordt. De vorm van predicatie die daarvoor nodig is is de eigenschapspredicatie van het eerste niveau: op die manier kan het tempus van de tweede-niveau vorm van predicatie worden gesimuleerd met behulp van eerste-niveau conceptueel materiaal. En perdurantisme is de gedachte dat een fysiek object door de tijd heen bestaat (voortduurt) analoog aan de wijze waarop het een bepaalde ruimte inneemt: in beide gevallen heeft zo'n object delen die zich op de relevante (temporele of ruimtelijke) locaties bevinden. Zoals een tafel zich over een bepaalde plaats uitstrekt omdat hij ruimtelijke delen heeft (poten en een blad), zo strekt diezelfde tafel zich over een bepaalde tijd uit omdat hij *temporele delen* heeft die gezamenlijk reiken vanaf zijn ontstaan tot zijn vergaan. Perdurantisme beschouw ik als een manier om invulling te geven aan B-theorie: de temporele delen van dingen hebben hun eigenschappen onafhankelijk van de tijd waarop ze bestaan, en daarom kunnen zij de brug vormen waarlangs de gehelen waar het delen van zijn hun eigenschappen op bepaalde tijdstippen kunnen hebben zonder dat daarvoor een temporeel gedifferentiëerde vorm van predicatie nodig is.

Fysieke objecten zijn, in deze eerste-niveau simulatie, te definiëren als de materiële inhoud van willekeurige stukjes ruimte-tijd. Er is een object dat de ruimte-tijd inneemt die ik in mijn gehele leven inneem, maar ook een object dat de ruimte-tijd

inneemt die ik inneem plus de ruimte-tijd die de Maan gedurende het Pleistoceen innam. Wij mensen beperken ons weliswaar tot objecten die een zekere stabiliteit hebben en die wij interessant vinden, maar dat maakt voor hun status als fysiek object verder niets uit.

Het beeld dat zodoende ontstaat is dat van een statische, vierdimensionale werkelijkheid (die zich zowel in tijd als in ruimte uitstrekt). Wat zich op de een of andere ruimtelijk-temporele locatie bevindt staat volledig los van wat zich elders bevindt: het is een samenhangsloos geheel van individuele verschijnselen.[400] Dat is het beroemde beeld dat Hume van de werkelijkheid schetste: slechts individuele verschijnselen, geen samenhang of noodzakelijke verbindingen daartussen. Alles wat wij aan samenhang menen te zien berust slechts op projectie van onze verwachtingen, die wij koesteren doordat we bepaalde toevallige regelmatigheden in het patroon van individuele verschijnselen tot nu toe hebben ontdekt.[401]

Wat de Humeaan motiveert is skepsis over conceptuele verbanden in het algemeen (een typisch verschijnsel voor het Moderne beeld). Omdat hij zulke conceptuele verbanden niet in de werkelijkheid meent te kunnen vinden (hoe zien ze er dan uit?) verbant hij ze naar het subjectieve domein (of reduceert hij ze tot wat er wél in de werkelijkheid te vinden is, naar zijn mening). Strikt genomen voert deze houding tot metafysisch anti-realisme of skepticisme. Want zelfs Lewis' simulatie van het begrip *fysiek object* in termen van ruimtelijke en temporele extensie is op zichzelf een begrip met bijbehorende conceptuele verbanden—ook al zijn die heel mager. Een minder strikte Humeaan, zoals Lewis, zal dus oogluikend willen toestaan dat dergelijke conceptueel dunne verbanden iets met de werkelijkheid te maken mogen hebben: dat maakt dat hij zich wat metafysica betreft tot het eerste conceptuele niveau wil beperken, en op zoek gaat naar eerste-niveau simulaties van alle hogere-niveau fenomenen.

Hoofdstuk 6: Het Tweede Conceptuele Niveau

Er bestaat een groeiende belangstelling voor een anti-Humeaanse benadering van centrale metafysische thema's. In dit hoofdstuk, dat verhoudingsgewijs het grootste deel van dit proefschrift beslaat, bespreek ik zo'n benadering voor oorzakelijkheid en voor tijd—twee fundamentele begrippen van het tweede conceptuele niveau. In

[400] Lewis vat zijn metafysische programma samen onder de slogan 'Humean supervenience': de gedachte is dat alles terug te voeren is op patronen in het geschetste samenhangsloze geheel van individuele verschijnselen [zie Lewis 1986b, pp. ix–x].

[401] Een Humeaanse kijk op natuurwetten en oorzakelijkheid bestaat er dan ook hoofdzakelijk in die twee begrippen terug te voeren op pure regelmatigheden in het Humeaanse mozaïek.

deze samenvatting bespreek ik enkele wezenlijke resultaten van mijn onderzoek op deze twee gebieden.

Oorzakelijkheid

Er bestaat een sterke Humeaanse traditie als het gaat om oorzaak en gevolg. De gedachte is dat er *in werkelijkheid* niets meer is dan opeenvolging van gebeurtenissen, en dat het benoemen van sommige opeenvolgingen als oorzaak en gevolg op de een of andere wijze terug te voeren moet zijn op regelmatigheden in het gehele Humeaanse mozaïek. Lewis verdedigt een geraffineerde versie van dit idee.[402]

Het tweede-niveau begrip van oorzakelijkheid daarentegen beschouwt de oorzaak als de reden voor het bestaan van het gevolg: de oorzaak *produceert* het gevolg; dat is het primaire begrip. Er kan dus geen sprake van zijn dat het gehele Humeaanse mozaïek gegeven is en de relatie van oorzakelijkheid op basis daarvan bepaald wordt—hoe de gebeurtenissen zich ontvouwen hangt juist af van wat wat zal veroorzaken.

Een beroemde tegenstander van Lewis, wat betreft oorzakelijkheid, is David Armstrong. Hij beschouwt oorzakelijkheid als een gegeven, en dan met name concrete, individuele gevallen van oorzakelijkheid ('singular causation'). Zijn lievelingsvoorbeeld is de druk die we op ons lichaam kunnen voelen: daar, zo Armstrong, nemen we oorzakelijkheid direct waar. Armstrong houdt echter wel vast aan de idee van een Humeaans mozaïek van gebeurtenissen (hij noemt het het vierdimensionele decor, 'the four-dimensional scenery'), maar meent dat hij het Humeaanse aspect ervan kan uitvlakken door er dergelijke niet-gereduceerde, individuele causale verbindingen tussen gebeurtenissen aan toe te voegen. In zijn ogen worden de regelmatigheden waarmee zulke causale verbindingen optreden verklaard door natuurwetten (die hij als noodzakelijke relaties tussen universalia ziet).

Armstrong is een interessant geval voor mijn onderzoek, omdat hij dus een sterk anti-Humeaanse kijk op oorzakelijkheid verbindt met een overigens eerste-niveau metafysisch plaatje. De kritiek die Lewis op Armstrong heeft komt dan ook in wezen neer op de klacht: je *zegt* wel dat je het hebt over 'individuele causatie' en over 'noodzakelijke relaties', maar wat *maakt* dan dat de betreffende relaties die naam verdienen? Het probleem voor Armstrong is dat zijn metafysica hem helemaal

[402] Lewis hangt een 'contrafeitelijke' (counterfactual) benadering van oorzakelijkheid aan, en baseert zijn notie van oorzakelijkheid dus niet alleen op regelmatigheden in het Humeaanse mozaïek maar ook op wat er in mogelijke alternatieve mozaïeken gebeurt. In deze samenvatting laat ik deze interessante complicatie buiten beschouwing.

niet toestaat om het begrip van causatie dat hij met zijn voorbeeld (druk op ons lichaam) illustreert serieus te nemen. Druk die je op je lichaam kunt ervaren is een sprekend voorbeeld van causatie als productie: de interactie tussen datgene wat druk uitoefent en je lichaam is wat de gevoelde druk produceert. De logica van dat soort oorzakelijkheid is niet die van een relatie op een gegeven vierdimensioneel decor van gebeurtenissen, want het vereist dat het ene relatum het andere produceert, terwijl de (eerste-niveau) soort relaties waar Armstrong mee werkt het reeds bestaan van de relata vooronderstelt.

Wat er nodig is om causatie als productie serieus te nemen is dus een radicaal andere kijk. Die is ook beschikbaar in de hedendaagse literatuur, namelijk in de vorm van een metafysica van vermogens ('powers metaphysics'). Dat is een sterkere vorm van anti-Humeanisme dan die van Armstrong: in plaats van de basale, ongeanalyseerde causale relaties van Armstrong gaat een metafysische theorie van oorzakelijkheid op basis van vermogens uit van een nieuw soort eigenschap, een vermogen dus, dat in een causale werking resulteert als aan bepaalde voorwaarden wordt voldaan. Merk op dat het begrip *vermogen* een begrip van het tweede conceptuele niveau is. Typische voorbeelden zijn disposities zoals oplosbaarheid en ontvlambaarheid: onder de juiste omstandigheden zal een oplosbare (of ontvlambare) substantie daadwerkelijk oplossen (of ontbranden). Vaak zullen in zulke situaties allerlei interventies mogelijk zijn die voorkomen dat het verwachte effect optreedt—maar het relevante causale proces is in zulke gevallen wel in gang gezet. Bijvoorbeeld: de substantie *was aan het oplossen*, maar *is niet volledig opgelost*. Dat is de relevantie van het aspect waarvan ik opmerkte dat het een rol speelt in de tweede-niveau vorm van predicatie (naast het tempus, waarover hieronder meer).

De Humeaan zal zulke eigenschappen willen reduceren, ten eerste omdat ze empirisch niet waarneembaar zijn (hoe ziet de oplosbaarheid van zout eruit?) en ten tweede omdat ze niet in een Humeaans mozaïek passen—ze leiden tot verdachte, niet-waarneembare verbindingen tussen afzonderlijke gebeurtenissen in de werkelijkheid. Vanuit een tweede-niveau perspectief is dat juist wat te verwachten is—dat is wat oorzakelijkheid *is*; de oorzaak is de reden voor het bestaan van het effect.

In plaats van het Humeaanse, statische idee van een over de gehele ruimte en tijd heen gegeven werkelijkheid opgebouwd uit afzonderlijke, samenhangsloze feiten (een Humeaans mozaïek), komen we zo tot een echt tweede-niveau plaatje van een werkelijkheid *op dit moment*, die zich dankzij de causale processen die zich aan het ontplooien zijn verandert en ontwikkelt. Dat brengt ons naar het tweede onderwerp van dit hoofdstuk: tijd.

Tijd

De metafysische discussie over tijd is complex, en kent bijzonder veel onderscheidingen. Drie daarvan noemde ik hierboven al, in onze samenvatting van hoofdstuk 5 (zie p. 372 hierboven): die tussen eternalisten en presentisten, die tussen B-theoretici en A-theoretici, en die tussen perdurantisten en endurantisten. Eerste-niveau denkers neigen veelal naar een eternalistisch, B-theoretisch, en perdurantistisch begrip van tijd. Net als in het geval van oorzakelijkheid, bestaat er ook hier een oppositie tegen een dergelijke visie—een oppositie die zich meer of minder specifiek richt op één of meer van de drie genoemde elementen. En, weer net als in het geval van oorzakelijkheid, valt op dat metafysici die zich op slechts één van de genoemde dimensies van een eerste-niveau visie op tijd willen distantiëren, maar overigens een eerste-niveau metafysisch verhaal aanhangen, zonder uitzondering falen. Het is zelfs mogelijk om op een eerste-niveau basis toch zowel presentist, A-theoreticus als endurantist te zijn. Dat maakt het lastig om een positieve visie op tijd als een begrip van het tweede conceptuele niveau te ontwikkelen.

Neem, om te beginnen, endurantisme: in tegenstelling tot de perdurantist, die zegt dat fysieke objecten door de tijd heen bestaan doordat ze zich met hun verschillende temporele delen op verschillende temporele locaties bevinden, wil een endurantist zeggen dat fysieke objecten door de tijd heen bestaan doordat ze zich *in hun geheel* door de tijd heen bewegen. Maar deze vrij voor de hand liggende gedachte kan nog steeds vanuit een eerste-niveau denkkader ingevuld worden, en geeft zodoende aanleiding tot een alternatieve theorie over het voortduren van objecten in de tijd—bijvoorbeeld door te zeggen dat de tijdsgebonden eigenschappen van objecten eigenlijk relaties tussen die objecten en tijdstippen zijn. Ter illustratie: dat de banaan gisteren groen was en vandaag geel is, is waar omdat de eigenschappen 'groen' en 'geel' relaties tot tijdstippen zijn. De banaan staat in de groen-relatie tot het tijdstip t_1 (gisteren) en in de geel-relatie tot het tijdstip t_2 (vandaag). Merk op dat wat eerst eigenschaps-toeschrijvingen waren middels een tempus (tweede-niveau predicatie) daarmee omgevormd wordt tot eigenschaps-toeschrijvingen middels de atemporele eerste-niveau vorm van predicatie. Deze vorm van endurantisme geeft dus een alternatieve invulling aan een B-theoretische benadering van tijd.[403] Merkwaardi-

[403] Het volgende kan wellicht verduidelijken. Als je een temporele uitspraak zoals 'de banaan is groen' formaliseert in standaard predicatenlogica dan krijg je bijvoorbeeld Gb (waar G voor groen staat, en b voor de banaan in kwestie). Het tempus valt bij zo'n formalisatie weg, en daarmee is de betrekking tot de tijd verdwenen. De perdurantist herintroduceert die betrekking door het object temporeel te kwalificeren: Gb_t. De endurantist die ik hier beschrijf herintroduceert die betrekking door de eigenschap temporeel te kwalificeren: $G(b, t)$ (of $G_t b$).

gerwijze blijken fysieke objecten volgens deze versie van endurantisme eerste-niveau entiteiten *buiten* de tijd te zijn.

Om tot een vorm van endurantisme te komen die zich fundamenteler van perdurantisme onderscheidt lijkt het dus zinvol om ook B-theorie te verwerpen. Dat brengt ons echter niet direct verder, want ook A-theorie laat versies toe die zich eigenlijk niet wezenlijk van B-theorie onderscheiden. Neem de beroemde temporele logica van Arthur Prior, die een fervent tegenstander van B-theorie was—volgens die logica kunnen we het tempus begrijpen als een soort modale operator op uitspraken in de tegenwoordige tijd. Zodoende kunnen we van een dergelijke uitspraak, zoals 'het regent' (p), nieuwe uitspraken maken, zoals 'het was het geval dat het regent' ($\mathbf{P}p$) en 'het zal het geval zijn dat het regent' ($\mathbf{F}p$). Op het eerste gezicht lijkt dat een goede manier om de tweede-niveau vorm van predicatie te formaliseren. Maar schijn kan bedriegen: met hetzelfde gemak kunnen we namelijk de basale uitspraken in de tegenwoordige tijd, zoals p, beschouwen als atemporeel en dus eerste-niveau. De idee daarbij is dat er een specifiek moment bestaat op de tijdslijn—het heden—waar dergelijke atemporele uitspraken op gesitueerd zijn. De operatoren **P** en **F** verschuiven dan het relevante tijdspunt over de tijdslijn. Het resultaat is een simulatie van het tempus met behulp van puur eerste-niveau middelen.[404]

Het probleem ligt dus in de voorstelling van tijd als een soort extra dimensie aan de werkelijkheid, analoog aan de ruimtelijke dimensies. Die voorstelling komt tot uitdrukking in de gedachte dat het temporele aspect van objecten, eigenschappen, of waarheden erin bestaat dat ze op een of andere manier een positie op die dimensie innemen. Nu zou je denken dat dit precies eternalisme is, en dat we dus eindelijk van de besproken eerste-niveau simulaties van tijd naar het echte, tweede-niveau begrip van tijd komen als we eternalisme verwerpen en dus presentisme aanhangen. Maar hier schuilt alweer een addertje onder het gras: er bestaat een versie van presentisme die toch nog aan de basale eternalistische gedachte vasthoudt. Die versie, die ik *negatief presentisme* noem, ziet er als volgt uit. De eternalist zegt dat de werkelijkheid zich over de gehele tijdslijn uitstrekt, dat ook toekomstige en verleden dingen en waarheden onderdeel ervan uitmaken (hoe we die verder ook begrijpen). De negatieve presentist nu ontkent dat en zegt dat de werkelijkheid *uitsluitend* het tegenwoordige moment omvat.

Het probleem is dat dit niets verandert aan de wijze waarop de eternalist het temporele aspect van objecten, eigenschappen of waarheden begrijpt: dat wordt nog

[404] Deze zienswijze lijkt op de reductieve benadering van modaliteit die Lewis aanhangt: daar zijn de basale uitspraken gesitueerd in de actuele wereld, terwijl de modale operatoren $\Box$ en $\Diamond$ ons naar andere mogelijke werelden verhuizen.

steeds in termen van een locatie op de tijdslijn begrepen. Het enige dat verandert is dat de negatieve presentist het bestaan van andere temporele locaties dan het heden ontkent. En daarmee raakt hij van de regen in de drup: als er zelfs al geen tijdslijn meer is, waarom zouden we het ene gegeven moment dan nog als een *moment in de tijd* beschouwen?

De uitweg bestaat erin het hele idee van temporele locaties en een tijdslijn, van de tijd als een dimensie aan de werkelijkheid, te verwerpen. Dat is dan ook wat ik voorstel. Preciezer gezegd stel ik voor om dergelijke voorstellingen als analogieën te begrijpen, die maar een beperkte geldigheid hebben, en die niet het fundament van de metafysica van tijd uitmaken. In plaats daarvan komt een begrip van tijd dat gegrond is in de tweede-niveau vorm van predicatie: tijd is de vorm van bestaan van fysieke objecten, met hun toestanden, vermogens, en activiteiten. Het voortschrijden van de tijd *is* de activiteit van de dingen die bestaan—tijd en oorzakelijkheid vormen zo twee zijden van dezelfde tweede-niveau medaille. Dit leidt inderdaad tot een begrip van tijd waar presentisme, A-theorie en endurantisme wezenlijk onderdeel van uitmaken, maar dan niet als posities waar je eternalisme, B-theorie en perdurantisme tegenover kunt plaatsen. Het punt is dat de hele discussie over tijd, op alle drie de besproken dimensies, *eigenlijk* een dispuut is tussen denkers die een eerste-niveau simulatie van tijd verdedigen en denkers die pogingen doen om daar bovenuit te stijgen en het tweede-niveau begrip van tijd te vatten. De sterk reductionistische tendens in de analytische wijsbegeerte heeft het denken in termen van eerste-niveau simulaties bijkans tot gewoonte gemaakt—dat blijkt wel uit mijn vaststelling dat veel van de schermutselingen op het slagveld van de metafysiche discussie over tijd uiteindelijk slechts onenigheden tussen verschillende eerste-niveau simulaties zijn.[405]

Hoofdstuk 7: Het Derde Conceptuele Niveau

Michael Thompsons uiteenzetting over leven, levensvormen, en de typische (derde-niveau) oordeelsvormen die we in ons begrip daarvan toepassen, vormde een belangrijke inspiratiebron voor mijn onderzoek op die gebieden [zie M. Thompson 2008, deel I]. De belangrijkste observatie die ik van Thompson overneem en in dit hoofdstuk verder uitbouw is de volgende: elke poging om tot een reductieve definitie van het begrip *leven* te komen faalt, oftewel doordat ze de plank misslaat, oftewel doordat ze het begrip ongemerkt inbouwt en daarmee circulair is. Neem bijvoor-

[405] In deze samenvatting laat ik enkele vraagstukken over de open toekomst en over de verhouding tussen mijn resultaten en de relativiteitstheorieën van de natuurkunde, die ik in §6.2.6 kort bespreek, buiten beschouwing.

beeld de idee dat levende wezens een heel bepaalde vorm van *organisatie* kennen die buiten de levende natuur niet voorkomt. Dat is natuurlijk waar, maar op het eerste gezicht ook circulair: het is immers een specifiek *biologisch* soort organisatie die het leven kenmerkt. Als het hier daarentegen echt om een reductieve definitie gaat, dan is dus een begrip van organisatie bedoeld dat ons *niet* impliciet terugverwijst naar het leven zelf. Maar dan is al snel duidelijk dat zo'n begrip ook van toepassing gaat zijn in situaties waar er van leven geen sprake is (zelfs als de definitie zo geraffineerd zou zijn dat zulke situaties in werkelijkheid nooit voorkomen). En met andere eigenschappen van het leven die vaak als kenmerkend worden beschouwd (zoals reproductie, homeostase, zelf-beweging, en ontwikkeling) staat het bij nadere beschouwing niet anders. Het leven vormt een categorie op zich.

Toch bestaat er een zekere consensus onder hedendaagse filosofen van de biologie over twee gedachten. De eerste is dat het leven in een bepaalde zin terug te voeren moet zijn op de tweede-niveau substanties en processen die eraan ten grondslag liggen, en die door bijvoorbeeld de moleculaire biologie tot in verbazend detail zijn onderzocht en beschreven. De tweede is de gedachte dat met de opkomst van de evolutietheorie de idee van levensvormen zoals ik die verdedig onder behoorlijke druk is komen te staan: als soorten evolueren, en dus gradueel door de tijd heen veranderen in andere soorten, hoe kunnen ze dan het soort essenties hebben dat ik voorsta? Wat ik naar aanleiding van deze beide skeptische gedachten over het derde conceptuele niveau te zeggen heb vat ik hieronder kort samen.

De voorstelling dat het leven uiteindelijk mechanistisch begrepen dient te worden is niets anders dan het dogma waar tweede-niveau metafysica op rust—dat de werkelijkheid *achter* de levende natuur (de werkelijkheid-op-zich) uiteindelijk in tweede-niveau termen begrepen dient te worden, en dat het derde conceptuele niveau slechts van toepassing is op de werkelijkheid-voor-ons.

In §7.3 bespreek ik, bij wijze van voorbeeld, de uitgewerkte tweede-niveau theorie van biologische functies van Ruth Millikan, en concludeer dat deze niet slaagt: het leven wordt, in de handen van Millikan, een vorm van contingente stabiliteit, waarvan duidelijk is dat die ook buiten het bereik van het leven kan bestaan. Millikan's reductieve theorie is interessant omdat ze daarnaast wel handig gebruik maakt van derde-niveau begrippen voor haar visie op thema's uit de taalfilosofie en de filosofie van de geest (zoals betekenis, taal, perceptie, en dergelijken): ze meent dat te mogen doen *juist omdat* ze die derde-niveau begrippen voorziet van een 'niet-normatieve, niet-mysterieuze' (lees: tweede-niveau) analyse.[406]

[406] Het citaat komt uit [Millikan 1984, p. 17]—op p. 287 citeer ik de betreffende passage vollediger.

Hoe moeten we dan de verhouding tussen het tweede- en het derde-niveau aspect van de werkelijkheid begrijpen, als we een dergelijke reductionistische visie van de hand wijzen? Hier zijn twee begrippen van belang: het begrip *formele oorzakelijkheid* en het begrip *materiaal*. Formele oorzakelijkheid speelt niet alleen op het derde conceptuele niveau een rol, maar ook al op het tweede conceptuele niveau. Neem, om een handzaam voorbeeld te gebruiken, een electron: als het electron in een nevelkamer geschoten wordt waar een bepaald magnetisch veld in heerst, dan zal het electron zich houden aan de regels die het electron-zijn hem opleggen—aan de conceptuele waarheden die uit het begrip *electron* volgen. Voor zover die regels diverse opties openlaten bestaan er diverse mogelijke routes die het electron zou kunnen afleggen; voor zover die regels maar één optie openlaten ligt de route van het electron vast (afgezien van mogelijke verstorende factoren, zoals een experimentator die het magnetische veld in kwestie manipuleert). Het spectrum aan mogelijkheden dat het electron heeft, in dit specifieke geval, is heel bepaald—en die vorm van bepaling noem ik, ruwweg aansluitend bij Aristoteles, formele veroorzaking.

In een derde-niveau context werkt formele veroorzaking net zo: de conceptuele waarheden die met het begrip *paardenkastanje* verbonden zijn bepalen wat er wel en niet mogelijk is voor de individuele paardenkastanje die buiten op het plein staat, gegeven de omstandigheden waarin hij zich bevindt. De paardenkastanje houdt zich zogezegd aan zijn eigen regels, net als een electron dat doet. Een belangrijk verschil echter is dat een electron een stukje materie *is*, terwijl de materie waar de paardenkastanje uit bestaat door hem als *materiaal* gebruikt wordt. En daarmee zijn we bij het tweede begrip aangeland dat ik hier kort wil introduceren.

Het materiaal waar een paardenkastanje uit bestaat laat zich ook bestuderen en beschrijven vanuit een tweede-niveau gezichtspunt. Van alle soorten materie die je in een paardenkastanje kunt vinden kun je het kookpunt, het smeltpunt, de oplosbaarheid in water, en allerlei andere meer of minder interessante chemische en fysische eigenschappen vaststellen. Maar de paardenkastanje maakt gebruik van die eigenschappen op een zodanige wijze dat de materie in kwestie het leven van de paardenkastanje ten goede komt. Dat wordt mooi duidelijk wanneer je typische biochemische processen bekijkt, zoals fotosynthese: dat het lichtgevoelige chlorophyl in de bladeren van de paardenkastanje een electron afgeeft bij binnenkomst van licht van de juiste golflengte, en daarmee een kettingreactie in gang zet die uitmondt in de synthese van glucose, is geen toeval, ook al is het maar één van de vele, vele chemische mogelijkheden die dat soort molecuul kent—*deze* eigenschap van chlorophyl is waar de paardenkastanje gebruik van maakt, en dat is dus het *doel*

dat chlorophyl dient (daarom is fotosynthese een teleologisch levensproces, dat in de samenhang van het leven van—in dit geval—een paardenkastanje staat, en niet slechts een tweede-niveau causaal proces).

De paardenkastanje kan alleen maar op deze wijze sturend ingrijpen in hoe zijn materiaal zich gedraagt als dat materiaal daar de ruimte voor laat. Daarom veronderstellen levende wezens *indeterminisme* op het tweede conceptuele niveau. We moeten, om met Elizabeth Anscombe te spreken, de natuurwetten die het materiaal beheersen beschouwen als de regels van een schaakspel—binnen die regels zijn, voor de levende wezens, vele zetten mogelijk, zonder dat de natuurwetten geschonden hoeven te worden.[407] Het is, uiteraard, een gedurfde stelling, dat organismes zoals een paardenkastanje hun onderliggende materiaal beïnvloeden op manieren die verder gaan dan wat door de aard van dat materiaal zelf bepaald wordt. Vanuit de gedachte van formele oorzakelijkheid is deze stelling minder gedurfd dan het lijkt; het is eenvoudigweg wat je krijgt als je het derde conceptuele niveau serieus neemt en niet ten prooi valt aan het dogma van tweede-niveau metafysica.

Dan rest mij nog de verhouding tot de evolutietheorie. Hier kan ik kort zijn: welbeschouwd is er helemaal geen spanning tussen de idee van evolutie en de idee van essenties in de vorm van levensvormen. Het is van meet af aan duidelijk dat levensvormen zich op velerlei wijze in de hun beschikbare materie kunnen uitdrukken, al naar gelang de omstandigheden en mogelijkheden ter plaatse. Levende wezens zijn bijzonder dynamisch en veelzijdig. Evolutie is een proces waarbij een door omstandigheden ontstane afwijkende instantie van een bestaande levensvorm aanleiding geeft voor de instantiëring van een andere, nieuwe levensvorm (wellicht eveneens in een afwijkende verschijningsvorm). Het is dus wel degelijk denkbaar dat de ene soort uit de andere voortkomt via een gradueel evolutionair proces, zonder daarmee de idee van levensvormen op te geven.

Afsluitende Opmerkingen

Het Aristotelische beeld vormt de basis voor een metafysica die op allerlei vlakken verschilt van veel van wat je in de hedendaagse metafysische discussie kunt vinden—die is immers, zoals ik in hoofdstuk 1 beargumenteerd heb, grotendeels gebaseerd op het Moderne beeld. Het meest wezenlijke verschil ligt in de wijze waarop metafysica begrepen wordt: aanhangers van het Moderne beeld gaan ervan uit dat de

[407] Zie Anscombe [1971, p. 143]. De passage waarin Anscombe deze vergelijking maakt citeer ik op p. 266.

begrippen die we voor ons inzicht in de werkelijkheid-voor-ons gebruiken op zichzelf nog niets zeggen over hoe onze theorie van de werkelijkheid-op-zich eruit moet zien, en beschouwen een reductieve visie op (delen van) de werkelijkheid-voor-ons daarom als de juiste route om tot een conceptueel zuinige en daarom betere theorie van de werkelijkheid-op-zich te komen. (Tenminste, voor zover ze metafysisch realisten zijn—metafysisch sceptici menen dat we nooit tot enig betrouwbaar resultaat kunnen komen in onze zoektocht daarnaar, terwijl metafysisch anti-realisten de hele zoektocht naar een dergelijke theorie verwerpen). Zulke metafysici zullen dus, afhankelijk van hun voorkeur, voor eerste-niveau metafysica of tweede-niveau metafysica kiezen.

Mijn onderzoek mondt uit in twee resultaten. Het eerste resultaat is dat de benadering van metafysica die in het Aristotelische beeld vervat ligt te prefereren is boven de heersende, op het Moderne beeld gebaseerde benadering. Het tweede resultaat is dat een vruchtbare manier van invulling geven aan een dergelijke Aristotelische metafysica begint bij een nauwgezet onderzoek naar conceptuele clusters die zich rondom verschillende vormen van predicatie groeperen. Ik heb een begin gemaakt met de uitwerking van deze benadering voor de besproken drie conceptuele niveaus, en tevens laten zien dat daarmee licht geworpen kan worden op allerlei fundamentele kwesties die in het tegenwoordige metafysische debat besproken worden.

* * *

Curriculum Vitae

Jesse Merijn Mulder was born on October 28th 1982 in Alkmaar, the Netherlands. He graduated with bachelor degrees in liberal arts and sciences and in philosophy from Utrecht University in 2008. He obtained a master degree (*cum laude*) in philosophy at Utrecht University in 2010. In 2009 he was a visiting scholar at the University of Connecticut for the duration of one semester. He has given courses, published articles, and given talks on topics in metaphysics, logic, philosophy of science, and philosophy of mind. His research interests include, in particular, essentialism, conceptual truth, philosophical methodology, (anti-)reductionism, consciousness, and free will. He happily lives in Bilthoven with his wife, Anne Geels, and their two daughters, Ronja and Momo.

* * *

Quaestiones Infinitae

PUBLICATIONS OF THE DEPARTMENT OF PHILOSOPHY AND RELIGIOUS STUDIES

VOLUME 21 D. VAN DALEN, *Torens en Fundamenten* (valedictory lecture), 1997.
VOLUME 22 J.A. BERGSTRA, W.J. FOKKINK, W.M.T. MENNEN, S.F.M. VAN VLIJMEN, *Spoorweglogica via EURIS*, 1997.
VOLUME 23 I.M. CROESE, *Simplicius on Continuous and Instantaneous Change* (dissertation), 1998.
VOLUME 24 M.J. HOLLENBERG, *Logic and Bisimulation* (dissertation), 1998.
VOLUME 25 C.H. LEIJENHORST, *Hobbes and the Aristotelians* (dissertation), 1998.
VOLUME 26 S.F.M. VAN VLIJMEN, *Algebraic Specification in Action* (dissertation), 1998.
VOLUME 27 M.F. VERWEIJ, *Preventive Medicine Between Obligation and Aspiration* (dissertation), 1998.
VOLUME 28 J.A. BERGSTRA, S.F.M. VAN VLIJMEN, *Theoretische Software-Engineering: kenmerken, faseringen en classificaties*, 1998.
VOLUME 29 A.G. WOUTERS, *Explanation Without A Cause* (dissertation), 1999.
VOLUME 30 M.M.S.K. SIE, *Responsibility, Blameworthy Action & Normative Disagreements* (dissertation), 1999.
VOLUME 31 M.S.P.R. VAN ATTEN, *Phenomenology of choice sequences* (dissertation), 1999.
VOLUME 32 V.N. STEBLETSOVA, *Algebras, Relations and Geometries (an equational perspective)* (dissertation), 2000.
VOLUME 33 A. VISSER, *Het Tekst Continuüm* (inaugural lecture), 2000.
VOLUME 34 H. ISHIGURO, *Can we speak about what cannot be said?* (public lecture), 2000.
VOLUME 35 W. HAAS, *Haltlosigkeit; Zwischen Sprache und Erfahrung* (dissertation), 2001.
VOLUME 36 R. POLI, *ALWIS: Ontology for knowledge engineers* (dissertation), 2001.
VOLUME 37 J. MANSFELD, *Platonische Briefschrijverij* (valedictory lecture), 2001.
VOLUME 37A E.J. BOS, *The Correspondence between Descartes and Henricus Regius* (dissertation), 2002.
VOLUME 38 M. VAN OTEGEM, *A Bibliography of the Works of Descartes (1637-1704)* (dissertation), 2002.
VOLUME 39 B.E.K.J. GOOSSENS, *Edmund Husserl: Einleitung in die Philosophie: Vorlesungen 1922/23* (dissertation), 2003.
VOLUME 40 H.J.M. BROEKHUIJSE, *Het einde van de sociaaldemocratie* (dissertation), 2002.
VOLUME 41 P. RAVALLI, *Husserls Phänomenologie der Intersubjektivität in den Göttinger Jahren: Eine kritisch-historische Darstellung* (dissertation), 2003.
VOLUME 42 B. ALMOND, *The Midas Touch: Ethics, Science and our Human Future* (inaugural lecture), 2003.
VOLUME 43 M. DÜWELL, *Morele kennis: over de mogelijkheden van toegepaste ethiek* (inaugural lecture), 2003.
VOLUME 44 R.D.A. HENDRIKS, *Metamathematics in Coq* (dissertation), 2003.
VOLUME 45 TH. VERBEEK, E.J. BOS, J.M.M. VAN DE VEN, *The Correspondence of René Descartes: 1643*, 2003.
VOLUME 46 J.J.C. KUIPER, *Ideas and Explorations: Brouwer's Road to Intuitionism* (dissertation), 2004.
VOLUME 47 C.M. BEKKER, *Rechtvaardigheid, Onpartijdigheid, Gender en Sociale Diversiteit; Feministische filosofen over recht doen aan vrouwen en hun onderlinge verschillen* (dissertation), 2004.
VOLUME 48 A.A. LONG, *Epictetus on understanding and managing emotions* (public lecture), 2004.
VOLUME 49 J.J. JOOSTEN, *Interpretability formalized* (dissertation), 2004.

VOLUME 50 J.G. SIJMONS, *Phänomenologie und Idealismus: Analyse der Struktur und Methode der Philosophie Rudolf Steiners* (dissertation), 2005.
VOLUME 51 J.H. HOOGSTAD, *Time tracks* (dissertation), 2005.
VOLUME 52 M.A. VAN DEN HOVEN, *A Claim for Reasonable Morality* (dissertation), 2006.
VOLUME 53 C. VERMEULEN, *René Descartes, Specimina philosophiae: Introduction and Critical Edition* (dissertation), 2007.
VOLUME 54 R.G. MILLIKAN, *Learning Language without having a theory of mind* (inaugural lecture), 2007.
VOLUME 55 R.J.G. CLAASSEN, *The Market's Place in the Provision of Goods* (dissertation), 2008.
VOLUME 56 H.J.S. BRUGGINK, *Equivalence of Reductions in Higher-Order Rewriting* (dissertation), 2008.
VOLUME 57 A. KALIS, *Failures of agency* (dissertation), 2009.
VOLUME 58 S. GRAUMANN, *Assistierte Freiheit* (dissertation), 2009.
VOLUME 59 M. AALDERINK, *Philosophy, Scientific Knowledge, and Concept Formation in Geulincx and Descartes* (dissertation), 2010.
VOLUME 60 I.M. CONRADIE, *Seneca in his cultural and literary context: Selected moral letters on the body* (dissertation), 2010.
VOLUME 61 C. VAN SIJL, *Stoic Philosophy and the Exegesis of Myth* (dissertation), 2010.
VOLUME 62 J.M.I.M. LEO, *The Logical Structure of Relations* (dissertation), 2010.
VOLUME 63 M.S.A. VAN HOUTE, *Seneca's theology in its philosophical context* (dissertation), 2010.
VOLUME 64 F.A. BAKKER, *Three Studies in Epicurean Cosmology* (dissertation), 2010.
VOLUME 65 T. FOSSEN, *Political legitimacy and the pragmatic turn* (dissertation), 2011.
VOLUME 66 T. VISAK, *Killing happy animals. Explorations in utilitarian ethics.* (dissertation), 2011.
VOLUME 67 A. JOOSSE, *Why we need others: Platonic and Stoic models of friendship and self-understanding* (dissertation), 2011.
VOLUME 68 N. M. NIJSINGH, *Expanding newborn screening programmes and strengthening informed consent* (dissertation), 2012.
VOLUME 69 R. PEELS, *Believing Responsibly: Intellectual Obligations and Doxastic Excuses* (dissertation), 2012.
VOLUME 70 S. LUTZ, *Criteria of Empirical Significance* (dissertation), 2012
VOLUME 70A G.H. BOS, *Agential Self-consciousness, beyond conscious agency* (dissertation), 2013.
VOLUME 71 F.E. KALDEWAIJ, *The animal in morality: Justifying duties to animals in Kantian moral philosophy* (dissertation), 2013.
VOLUME 72 R.O. BUNING, *Henricus Reneri (1593-1639): Descartes' Quartermaster in Aristotelian Territory* (dissertation), 2013.
VOLUME 73 I.S. LÖWISCH, *Genealogy Compostion in Response to Trauma: Gender and Memory in 1 Chronicles 1-9 and the Documentary Film 'My Life Part 2'* (dissertation), 2013.
VOLUME 74 A. EL KHAIRAT, *Contesting Boundaries: Satire in Contemporary Morocco* (dissertation), 2013.
VOLUME 75 A. KROM, *Not to be sneezed at. On the possibility of justifying infectious disease control by appealing to a mid-level harm principle* (dissertation), 2014.
VOLUME 76 Z. PALL, *Salafism in Lebanon: local and transnational resources* (dissertation), 2014.
VOLUME 77 D. WAHID, *Nurturing the Salafi Manhaj: A Study of Salafi Pesantrens in Contemporary Indonesia* (dissertation), 2014.
VOLUME 78 B.W.P VAN DEN BERG, *Speelruimte voor dialoog en verbeelding. Basisschoolleerlingen maken kennis met religieuze verhalen* (dissertation), 2014.
VOLUME 79 J.T. BERGHUIJS, *New Spirituality and Social Engagement* (dissertation), 2014.
VOLUME 80 A. WETTER, *Judging By Her. Reconfiguring Israel in Ruth, Esther and Judith* (dissertation), 2014.
VOLUME 81 J.M. MULDER, *Conceptual Realism. The Structure of Metaphysical Thought* (dissertation), 2014.

Quaestiones Infinitae

$\theta\pi$