EXPERIENCE AND BEYOND

The Outline of a Darwinian Metaphysics

JAN FAYE

Experience and Beyond

Jan Faye

Experience and Beyond

The Outline of a Darwinian Metaphysics

palgrave
macmillan

Jan Faye
Department of Media, Cognition, and Communication
University of Copenhagen
Copenhagen, DK-2300 Denmark

ISBN 978-3-319-31076-3 ISBN 978-3-319-31077-0 (eBook)
DOI 10.1007/978-3-319-31077-0

Library of Congress Control Number: 2016947274

Printed on acid-free paper

This Palgrave Macmillan imprint is published by Springer Nature
The registered company is Springer International Publishing AG Switzerland

Preface and Introduction

Thanks to W.V.O. Quine and a number of other thinkers, a fair proportion of today's philosophers realize that the sciences may help us to answer epistemological questions. Usually, it is not physics they have in mind but psychology and cognitive science. The degree to which these philosophers are able to commit themselves to the conclusions of contemporary psychology and cognitive science depends on their willingness to explain the normative side of epistemology in naturalistic terms. But we should remind ourselves that humans are first and foremost animals, no more unique than any other animal, although no doubt by reflection more species-centric than others. Thus, even among those philosophers who seize upon the results of psychology and cognitive science in their thinking, there is a tendency to forget our human origin and how our mental powers came into existence. Still, some early pioneers and an increasing number of knights of reason have picked up the gauntlet, arguing that a naturalized theory of knowledge should be armored in evolutionary epistemology before it turns to social epistemology.

Not long ago the philosophical landscape was very different. That philosophers such as Edmund Husserl and the phenomenological school believed that consciousness did not have its roots in nature, but nature had its foundation in our reflective consciousness cannot come as much of a surprise. According to this anti-naturalistic approach, the human mind is completely alienated from the natural order. However, the situation was not much different within the analytic traditions. Here the Platonic forms of language were still considered to be objects for philosophical contemplation. Ludwig Wittgenstein directly abjured naturalism and its entire works in *Tractatus*: "The Darwinian theory has no more to do with philosophy than has any other hypothesis of natural science" (4.1122).

With few exceptions it was not until around 1970 that some philosophers in the analytic tradition began to take evolution and cognitive science seriously as relevant for understanding human knowledge. Saying this, I do not want to underplay the fact that some form of naturalism has been quite common in many forms of American philosophy, more or less coeval with pragmatism. John Dewey was certainly deeply influenced by Darwin, as was Charles S. Peirce, as well as other minor pragmatists. In recent years a vanguard among philosophers has successfully begun to implement Darwinism into their theories of knowledge. This approach will grow even stronger in the future, I believe, alongside progress in evolutionary biology, animal behavioral studies, and cognitive science. I do not believe we are likely to see more armchair philosophy of the sort that has been so abundant in the past.

Yet, what is missing by even these advocates of naturalistic epistemology is the consequence that evolutionary epistemology might have for finding a *metaphysical* foundation of science. If our faculty of reason is a result of natural selection by the environment in which our ancestors lived, one may expect that our innate cognitive powers are more limited than we commonly think. Indeed, scientists have been able to formulate surprising insights into the world that go far beyond anything we can imagine solely from our immediately experience. But in order to get to this point, they have introduced very abstract theories whose predictive and explanatory success is no guarantee for representational success.

The main purpose of the present book is to explore the relationship between scientific knowledge and possible metaphysical interpretations in the light of our Darwinian heritage. Assuming a naturalistic stance from the outset, I ask what it has to teach us about the inference from scientific observation to metaphysics (the same inference with which Kant was concerned from his transcendental stance). Like Kant I want to set limits to this inference. So with "naturalist" replacing "transcendentalist," Kant and I are associates. The particular form of naturalism for which I am arguing (by example) is of course based on the inference from "Darwinian" evolutionary theory to what I hold to be the limits of scientifically informed metaphysics or, for that matter, of any metaphysics. After having delineated, in Chap. 1, the manifest image and Kant's disavowal of any metaphysics of the things-in-themselves, I turn to a discussion of naturalism and evolutionary epistemology in Chap. 2. Here I argue in favor of a weak version of evolutionary metaphysics by holding that natural selection has empowered our predecessors with awareness, memory, imagination, intentions, learning mechanisms, and focused actions and that these innate dispositions for reflective thinking eventually replaced natural selection as the obvious grounds for understanding human cognitive evolution.

Science, I believe, gives us knowledge about concrete things we cannot see with the naked eye (*pace* Bas van Fraassen). This view is something I argue in Chap. 3. But many other things, which a realist interpretation of scientific theories represents as real, according to another, more instrumental, interpretation, are merely a result of hypostatized concepts and therefore unreal. Instead, these concepts are to be considered as designed for the sole purpose of understanding concrete facts coherently and continuously. This is the main claim of Chap. 4. Thus, my aim is to reveal the scope and limits of metaphysical inquiries on the assumption that Darwin was right about human evolution, not only with respect to our physiology but also with respect to our cognitive apparatus. I call the outcome of this analysis *evolutionary naturalism*. Of course, this stance is not new. Knowledgeable readers would know that Roy Wood Sellars published a book with this exact title in 1922, and in spite of the fact that I was unaware of his book until quite recently, it is quite obvious that we aim at some of the same philosophical tasks although our emphases are somewhat different.[1]

In my version, evolutionary naturalism is a metaphysical stance holding that a philosophical interpretation of our knowledge of reality

[1] Sellars, Roy Wood (1922). This also brings up the "school" of the American Realists who flourished in the 1920s and 1930s. Many of these guys called themselves "naturalists" as well as "realists" and were enthusiastic about Darwin. Besides Sellars there was Arthur O. Lovejoy, Ralph Barton Perry, and of course George Santayana. They were more or less constantly interacting with the second-generation pragmatists such as C.I. Lewis.

should be grounded in our biological powers and therefore that evolution has constrained our notion of reality to what is accessible from the common-sense world of sensory appearances and bodily interactions. In this perspective scientific understanding will always be grounded in common experience, although it also develops advanced concepts based on experiments and the application of mathematics to experience that refer to objects not directly accessible to the human senses. Thus, evolutionary naturalism is a metaphysical stance that maintains the existence of a physical reality separated from human beings but avoids any claim that our knowledge of this reality goes beyond what can be related to human experience. Its metaphysical point of view falls somewhere between full blown realism and meager instrumentalism. This implies, not surprisingly, that the evolutionary naturalist partly shares views with pragmatists, contextualists, and ontological pluralists.

Darwin hit the nail on its head, I think, when he wrote in his diary: "He who understands baboons would do more towards metaphysics than Locke."[2] The human mind is not an unprepared *tabula rasa* on which reality writes information. Selection has given us a variety of cognitive schemata that help us to grasp sensory information and enable us to act and reason based on previous experiences. In some sense Kant was closer to the truth as he argued that our cognitive ability involves the categories of the understanding that are imposed on our sense impression. But he was too much enmeshed in his rationalist heritage in insisting on the necessity and universality of these allegedly a priori categories. The human species already inherited some forms of thinking from its animal predecessors to survive and reproduce in whatever environment it had to face. Nevertheless, when these cognitive schemata were selected they were not predestined in relation to either the organism or the environment. They were not necessarily even the best possible, but the best available. As the human species developed its capacity of language it eventually gained a capacity of describing the world according to these schemata. All later reflection and the use of language are adapted to obey these schemata.

Others have attempted to naturalize metaphysics. But the naturalization of philosophy has meant different things to different philosophers. The most

---

[2] Quoted from Ruse (2012), p. 137.

widespread approach is to use the same empirical methods as the sciences to justify philosophical claims. The difficulty is that many philosophical problems are not empirical issues at all and therefore cannot be determined experientially. Dropping Kant's transcendental approach does not imply an empiricist view of philosophy itself. A more groundbreaking approach is to adopt an ontology distilled from the sciences, and here physics is considered to be second to none. But this approach also has its pitfalls. Often such an attitude presupposes scientific realism, which is insufficiently grounded in physics itself. Also scientific realism tends to reductionism, which implies that the ontologies of biology, psychology, or sociology have to give way to the ontology of physics. As an alternative I shall offer an evolutionary ontology that I contend is more in accordance with common-sense realism. This ontology does not eschew the discoveries of physics, but it is not grounded in a reductionist scientific realism. Rather, it builds on the same natural instincts of realism in which common-sense realism arises. To be in line with what we know from evolutionary biology a realist of this opinion has to argue that much of the picture physics draws is human constructions, having no counterpart in a scientifically supported notion of reality. Common-sense realism is egalitarian; it incorporates the discoveries of sciences, but it also accords them equal treatment.

A critical approach toward any attempt to build a naturalized metaphysics based on evolution has recently been aired by James Ladyman and Don Ross in their defense of "physicism." The following quotation encapsulates their dismissal: "[P]roficiency in inferring the large-scale and small-scale structure of our immediate environment, or any features of parts of the universe distant from our ancestral stomping grounds, was of no relevance to our ancestors' reproductive fitness. Hence, there is no reason to imagine that our habitual intuitions and inferential responses are well designed for science or for metaphysics."[3] Instead, they endorse fundamental physics as metaphysically prior to the special sciences. Although Ladyman and Ross do not consider themselves to be scientific realists of the traditional ontological breed, but think of themselves as structural realists, they believe that people eventually have "learned to represent the world and reason mathematically—that is in a manner that enables us to abstract

[3] Landyman and Ross (2007), p. 2.

away from our familiar environment." Of course I fully agree that this ability has contributed enormously to the acquisition of scientific knowledge. Moreover, I do not disagree on the latter point. I also think that scientific explanations are far superior to habitual intuitions, for our intuitions change frequently partly due to the impact of scientific beliefs. But apart from these points of agreement I advocate the very view they attack. The problem is, I think, that a strong realist interpretation of the theoretical structure of science may reify or hypostasize what in fact really belongs to the false idea that scientific theories *represent* the world as it is in itself. I shall return to this issue in Chap. 4 as well as in Chap. 10. So my discussion is not about empirical knowledge *per se*, but about how far we can draw specific metaphysical conclusions based on the way scientific theories are presumed to represent the world.

It is a characteristic of any naturalized metaphysics that it rejects the possibility of a priori knowledge. I do not lament the loss. But Ladyman and Ross see metaphysics as a form of unification of science based on scientific explanation and justification. In contrast, I see it as partly an empirically underdetermined discipline and partly an examination of the interrelationships among the basic concepts by which we understand reality. Categories like truth, objects, events, space, time, and causation are all fundamental notions that all of the sciences presuppose and take for granted. These basic concepts are the issues of the discussion that takes place in Chaps. 5, 6, 7, 8, 9, and 10. I believe these categories are important for our understanding of what science does but cannot be fully explained by science itself. The provided analysis is revisable, not by empirical means, but by human reflection in case its claims are not in accord with our knowledge of human cognitive capacities and institutional norms and values. Metaphysics covers basic ontology, but it includes even more than that. It is a theory of *alleged* truth-makers; yet, this must include an exposition of how these alleged truth-makers become truth-makers. Metaphysics is the only discipline that delineates a possible story of the relationship between the world and our thinking about this world.

So I argue that evolutionary naturalists are metaphysical antirealists with respect to abstract concepts and 'things' that do not exist in empirical space and time. All those purported entities we cannot either directly see perceptually or indirectly detect experimentally are constructed entities in the sense that they exist only as human conceptions and not as natural entities. Evolutionary naturalism denies that abstract concepts are real entities and therefore that they can act as *real* truth-makers. In contrast to metaphysical realists, evolutionary naturalists restrict the number of real truth-makers to only those that observation can detect.

The idea behind the present book took form already during my stay at the Center for Philosophy of Science in Pittsburgh more than 20 years ago when I had completed my book on the Danish physicists Niels Bohr. In that book I characterized him an objective antirealist, i.e., a realist concerning elementary particles but an antirealist with respect to the quantum theory. In this respect I argued that he was very much like a pragmatized Kantian.[4] My original manuscript turned out to be much too large, too kaleidoscopic and unfocused, and therefore was never brought to a completion. I am happy for that, because my view today has a stronger Darwinian twist than it had in those days. This has given me a more empirical basis for defending my form of naturalism, which denies that scientific knowledge goes beyond what can be experienced directly or indirectly. However, the various themes of the manuscript survived, and some of them are treated in my other books. But in my hands the main topic of the relation between scientific knowledge and metaphysics remained untouched up to now.

Whom am I going to thank? Well, first and foremost the other visiting fellows at the Center, of whom many were from Germany. We all had time for a dinner or a drink late in the evening while discussing philosophical matters. I especially want to mention Thomas Uebel, with whom I shared an office, and Max Urchs, one of the many East Germans who in the German reunification were squeezed between the socialist tradition and the capitalist *Wissensdünkel*. Later, for some years he and I, together with Uwe Scheffler, arranged the Poel workshops on logic and philosophy of science, involving philosophers from countries around the Baltic Sea. Indeed, I also want to thank people around the Center such as John Earman, Adolf Grünbaum, Peter Machamer, Ted McGuire, John Norton,

[4] In spite of our slightly different views on Bohr, both Henry Folse (1994) and I agree that Bohr can be seen as a pragmatized Kantian.

Gerald Massey, and Nicholas Rescher for their hospitality and comments in discussions. Closer to home I ought to mention Finn Collin whose long-time criticism of the social constructivists is one I share. Many more have been of great inspiration, even some I may not remember today. Last, but not least, I must recognize Henry Folse who, as usual, stood by my manuscript with comments and linguistic corrections. He and I shared a common interest in Bohr's philosophy very early on, and when we met for the first time in Moscow at the 8th International Congress of Logic, Methodology, and Philosophy of Science in 1987, none of us could know that this encounter would bring us in such close contact over so many years. With these words let us move on.

Contents

1

Evolutionary Naturalism

Philosophers have obsessed over both the ontological question of what makes claims about the world true and the epistemological question of how we can justify believing that some claims are in fact true. Ordinary sensory perception provides most of our experience of the world, so the answer seems straightforward: facts about those things that we can see make our beliefs about them true. However, experience, in my use of the term, is the product of both bodily interaction and perceptual information. The regularity in our perception of things and our ability to intervene and manipulate what we can perceive assure us that beliefs based on sensory information are true. In spite of the fact that we sometimes misperceive things, evolution has given us strong confidence in the perceptual information provided by our senses because of the utility this information has for successful actions and life management.

But what about beliefs concerning things we cannot experience, say, historical events, future events, other minds, invisible entities, laws of nature, mathematical objects, or other entities posited by various metaphysical theories? The truth of all these beliefs seems to depend on different domains of reality, for which our senses are not biologically adapted. How can we, for instance, philosophically vindicate our belief that those

J. Faye, *Experience and Beyond*,
DOI 10.1007/978-3-319-31077-0_1

invisible entities and properties science posits are real? And, again, how can we philosophically justify believing that well-established theories science accepts are, within the conceptual scheme in which they are expressed, at least approximately true descriptions of some aspect of reality?

Wilfrid Sellars has addressed this problem by distinguishing between the *manifest image* of the man-in-the-world and the *scientific image* that has been arrived at through experimental investigation. The manifest image is a product of the conceptual framework in terms of which we observe and explain our everyday world. In this image the world consists of persons and things. Moreover, persons are different from things because people have thoughts and consider themselves to act for reasons. Thus, the manifest picture seems to correspond to the image of the world to which we are adapted by natural selection, whereas the scientific picture is the one that is hypothesized by our best physical theories. Sellars himself thought that the manifest image and the scientific image each constitute a unique frame of understanding that is prima facie mutually incomparable: "The scientific image presents itself as a *rival* image. From its point of view the manifest image on which it rests is an 'inadequate' but pragmatically useful likeness of a reality which first finds its adequate (in principle) likeness in the scientific image."[1] Hence, Sellars Junior emphasized the primacy of the scientific image in all cases where this image conflicts with the manifest image.[2] Nevertheless, he argued ultimately for a "synoptic vision," according to which the descriptive and explanatory resources of the scientific image are combined with the "language of community and individual intentions." The manifest image supplies the scientific image with a conceptual framework of persons, which "provide[s] the ambience of principles and standards (above all, those

[1] Sellars (1963), p. 20.

[2] Bas van Fraassen (1999) construes Sellars' position as if Sellars "argued that the two world pictures are in irreconcilable conflict, and that the infinitely superior scientific image must eventually displace the manifest image altogether" (p. 2). However, Sellars did not say anything as radical as that, but argued pretty unambiguously that "the conceptual framework of persons is not something that needs to be *reconciled with* the scientific image, but rather something to be *joined* to it" (p. 23). So even though Sellars perhaps considered the scientific imagine superior, he did not hold, as Fraassen accuses him of doing, that the manifest picture has to be replaced by the scientific image. Sellars took the two images to form a single stereoscopic picture.

which make meaningful discourse and rationality possible) within which we live our own individual lives."[3]

Sellars Junior neither claimed that the objects in the manifest image can be reduced to systems of invisible entities nor argued that the objects of the manifest world are fundamental and that the objects posited by science are mere abstractions constructed for prediction. Using Kant's terminology we can say he believed that the manifest image represents the world as it appears to us, whereas the scientific image in the long run represents things as they are in themselves. Also Sellars Senior, Roy Wood Sellars, had been occupied with finding "the inclusion of man in nature, an inclusion that would do justice to all his distinguishing characteristics."[4] Few philosophers before him had foreseen how this problem might be solved by seeing science as part of nature herself: "Science and philosophy are properties of man. To explain them, we must comprehend man's capacities and his place in the world."[5] For Sellars Senior the solution to the understanding of human beings and their capacities for science and philosophy lies in the explanation of man's appearance in nature. We find such an account in the view of evolutionary naturalism.

[3] Ibid., p. 40.

[4] Sellars (1922), p. 3.

[5] Sellars (1922), pp. 1–2. The context of quotation is "Philosophy like science is a human achievement, and so rests on man's capacities. Unlike science, philosophy is forced to consider those capacities and processes, which make it possible. It is for this reason that philosophy is necessarily so engrossed with man. Knowledge is a human affair, even though that which is known is distinct from the knower. But man is part of nature, and so these capacities and processes operative in science and philosophy must find their natural explanation. Intelligence must be given its locus and attachments. In order words, science and philosophy are properties of man. To explain them, we must comprehend man's capacities and his place in the world. The final problem of philosophy is to connect the fact and content of knowledge with its conditions. How does knowing occur in the kind of world that is actually known? Knowing is a fact and must be connected up with the world which sciences study. Thus a system of philosophy answering this question is the capstone of science."

1.1 The Manifest Image

In this work I attempt to formulate an informed stand on those problems touched upon by Sellars' distinction. What is real? To what extent does science represent the world as it really is? And does a separation between the manifest and scientific image make sense in relation to human evolution? I shall argue that, as it turns out, the distinction is not quite adequate. Evolution has given us the manifest image, the reality of common sense, but evolution also constrains how we can develop the scientific image. Science deals with both visible and invisible entities and correctly ascribes existence to both, but I also claim that in order to reconcile contemporary science with our biological heritage we must recognize that some of the basic categories of science are mere abstractions, i.e., constructions of our own minds. With respect to the issue of invisible entities, I hold that the reality of scientific objects need not be more problematic than that of everyday objects: From the point of view of an evolutionary epistemology, the challenge of proving the existence of things we cannot see is no different or more serious than the challenge posed by visible ones. The problem concerning the existence of scientific objects is simply part of the more general question of the extent to which we are warranted in believing in the reality of a mind-independent world.

With respect to the issue of whether science provides us with a true picture of the world, I suggest that we often describe and explain the world of concrete particulars in terms of abstract categories that we have somehow derived from experience and that therefore do not exist independently from the minds of scientists. We make such constructions in order to grasp our experience as of an organized and structured world existing in space and time. I also argue that we have developed these abstract categories *as identity conditions for external objects*. Not only the very forms of sensuous intuitions, space and time, but also the connection of phenomena through the category of causation and the various laws of nature derived therefrom are such abstract entities of just this sort. All function as various identity conditions for describing concrete particulars.

Humans need to create abstract concepts in order to be able to carry out reflective thinking, given the spatio-temporal limitations of the sensory input to which the progenitors of the human species were subjected. Our ancestors' sensory experience was confined to their immediate surroundings and was perspectival at the outset, so much, if not most, of our cognitive capacity rests on innate dispositions that facilitate our ability to learn in order to react and survive. But during the evolution of higher animals, and especially hominids, *Homo sapiens* also gained the capacity of abstract thinking, which started out as the ability to generalize types from tokens.[6] This ability then became strongly reinforced by the later evolution of human linguistic powers, suggested to have taken place between 200,000 and 100,000 years ago.[7] The ability to think abstractly turned out to be not only connected with the capability of generalization, but also with the capacity of contemplation, with planning and foreseeing a possible future based on past experiences.[8]

[6] There is a strong tendency in analytic philosophy to argue that conceptual and semantic (re)presentations go hand in hand. I think it is a mistake. Thoughts in higher animals seem to have been present long before our linguistic competence evolved. Wilfrid Sellars gave a sophisticated characterization of this tendency. In his (1963) he says about man's manifest image of himself-in-the world: "Its central theme is the idea that anything which can properly be called conceptual thinking can occur only within a framework of conceptual thinking in terms of which it can be criticized, supported, refuted, in short, evaluated. To be able to think is to be able to measure one's thoughts by standards of correctness, of relevance, of evidence. In this sense a diversified conceptual framework is a whole which, however sketchy, is prior to its parts, and cannot be construed as a coming together of parts which are already conceptual in character. The conclusion is difficult to avoid that the transition from pre-conceptual patterns of behaviour to conceptual thinking was a holistic one, a jump to a level of awareness which is irreducibly new, a jump which was the coming into being of man" (p. 6). In contrast, several studies indicate that nonhuman animals have thoughts, which include the use of different forms of nonlinguistic concepts. For instance, one comparative study conducted by Zentall *et al.* (2008) claims: "We suggest that several of the major varieties of conceptual classes claimed to be uniquely human are also exhibited by nonhuman animals. We present evidence for the formation of several sorts of conceptual stimulus classes by nonhuman animals: *perceptual classes* involving classification according to the shared attributes of objects, *associative classes* or functional equivalences in which stimuli form a class based on common associations, *relational classes*, in which the conceptual relationship between or among stimuli defines the class, and *relations between relations*, in which the conceptual (analogical) relationship is defined by the relation between classes of stimuli. We conclude that not only are nonhuman animals capable of acquiring a wide variety of concepts, but that the underlying processes that determine concept learning are also likely to be quite similar" (p.13).

[7] This is the estimate made by a recent study based on linguistic data, which is an estimate claimed to be consistent with fossil and genetic data. See Perreault and Matthew (2012).

[8] Abstract and contemplative thinking allows us to imagine possible scenarios for the future, but generally all it allows is expectations of a range of possible futures, not which of those futures will actually happen. The case of deterministic mechanistic prediction is relatively a very special case where we can predict *the* one and only possible future. That ability arrived only many millennia after the capacity for abstract thinking had evolved.

The introduction of language gave *Homo sapiens* a forceful tool to construct expressions that were independent of a reference to here and now, but based on the ability to imagine what could happen elsewhere at other times.[9]

Human evolution and natural selection have, I submit, strong implications for the metaphysical dispute between realists and antirealists. Until now this dispute has taken place without much notice of the possible constraints natural selection may put on our ability to pose and answer metaphysical questions. One holds a realist view with respect to some entity if one believes that this entity exists independently of any human categorizations and judgments. Being a realist is always relative to a class of entities; one does not have to be a realist with respect to every entity posited by successful scientific theories. Consequently, realists hold that statements of a certain class, whether they concern philosophical principles, mathematical objects, nature, history, other minds, or meanings, not only may be true, but also if true, they are true quite independently of our individual or collective preferences. Conversely, the antirealist is one who denies the cognitive independence of such entities and therefore denies that some of these types of statements, if not all of them, are true or can be held to be true irrespective of our ability to justify them. Therefore one may ask, how can we decide what is real and what is not if so much of our thinking involves abstractions and idealizations?

We are all realists concerning our everyday world because nature has supplied us with a realist instinct. Natural selection has molded human cognitive capacities such that sensory and behavioral information about our environment is automatically taken to be trustworthy and is not disbelieved unless other sensory or behavioral information overrules it. When other information 'overrules' my normal trust in my senses, it causes me to disbelieve the direct report of my senses, but it also leads me to 'correct,' not to discard that information. Consider the oar that looks bent where it intersects the water surface; it does not lead me to abandon my perception of the oar, but to correct it. Furthermore, it generally gives an explanation of why in this case the senses cannot be trusted. It is by

[9] Evolutionary biology is an empirical science, and if these sorts of claims are more than just speculative hypotheses, they must be justified by empirical evidence. In note 4 I mentioned Zentall et al. (2008). Further evidence can be found in Shettleworth (2010) who discusses the cognitive evolution in non-human animals.

explaining how the senses are deceived (a trick of perspective or different indices of refraction, etc.) that I can 'correct' for the deception. So selection is our guarantee that we have reliable and permanent knowledge about an independent environment that we do not control for the most part. It is plausible to suppose that natural selection has given, say, bees capacities that give them "reliable and permanent knowledge about an environment…" for the very same general reasons as for *Homo sapiens*. But because the goals and needs of bees' survival and reproduction are quite different from *Homo sapiens* (i.e., biologically we are quite different organisms), different sensory capacities have been selected, thereby making "bee knowledge"—and hence "bee reality" (the world-for-bees) very different from human knowledge and human reality. Yet, presumably the actual real world is the same one that both humans and bees inhabit. In some sense the challenge here is to explain how bee knowledge and bee reality can be as different from their human counterparts as (I think it is reasonable to suppose) they are and yet both be "of" the one-and-only same world we both inhabit. I see the flower and the bee sees the same flower, but what I see is not the same as what the bee sees. If those things that we experience did not exist, no suitable sensory modalities to accrue information about them would have developed in the long run. We are adapted to and by a mind-independent world. Our ability to both control parts of our environment and sometimes foresee or predict what we cannot control supports our realist instinct.

Most philosophers would believe that it would change the game entirely if the existence of invisible things hinged on our cognitive capacities. But we have no faculties other than our sensory capacities that reveal whether the thing we cannot perceive with our senses exists at all, much less that it does not exist, as we perceive it. This raises the question: If the flower is real, does it exist as the bee perceives it or as we perceive it, or as both or neither? The manifest image of bees is different from the manifest image of human beings. What is common to both is that natural selection has made bees and human beings sensitive to flowers. Thus, sensation, human or otherwise, *never* corresponds to an objective reality; it determines a reality as it appears to humans or as it appears to bees. Some would argue therefore that beyond the various appearances the flower objectively consists of invisible things that are different from how it appears to bees or human beings. Science aspires to tell us what these things are like.

Nevertheless, the realist about invisible entities must face a dilemma. She might be confronted with empirically underdetermined theories concerning invisible things. William H. Newton-Smith has aptly characterized the situation in the following way[10]: In this situation the realist cannot simultaneously satisfy the ontological or epistemological ingredients in her position. If she maintains the ontological ingredient, claiming that one theory is true and all others are false, then she cannot satisfy the epistemological ingredient because she possesses no means to determine which one is which. Newton-Smith calls such a view the *ignorant* response. It accepts the existence of inaccessible facts. If instead the realist attempts to satisfy the epistemological ingredients, arguing that empirically undecidable assertions do not involve any determinate facts that make them either true or false, then she cannot uphold the ontological ingredient, because if we adhere to the epistemological ingredient of realism, there are no inaccessible facts beyond our sensory experience. Newton-Smith calls this view the *arrogance* response.

1.2 The Scientific Image

The realist dilemma can also be used to approach the second issue. Do scientific theories provide knowledge about mind-independent laws governing an objective reality? Modern science makes extensive use of very unfamiliar concepts in its description of experimental and observational findings as well in its pursuit of purely theoretical insight and understanding. For example, in the last 40 years, physicists have tried to establish a consistent theory of everything, couched in an extraordinarily abstract formalism, according to which they claim to be able to explain all known forces in nature from the laws of "particles," so-called "superstrings," in

[10] Newton-Smith (1981), pp. 40–42. This would be a philosophical (i.e., in principle) objection only if these alleged facts are necessarily, universally inaccessible. If there is a possibility of some future way of accessing these facts, then we would find out which of the underdetermined theories (if any) is the true one, and the realist position all along would have been correct: one is true but we are ignorant of which.

an 11- or 24-dimensional space-time. Alas, some physicists have even attempted to postulate a space consisting of more than a thousand dimensions in return for a physical system with only one free variable. A pervasive opinion prevails among these theoreticians that, when all is said and done, the subatomic world is constructed from just this one kind of particle and hence that the existence of the entire macroscopic world is the effect of the behavior of these strings. A successful superstring theory would have to explain why we experience only four space-time dimensions instead of 11 (or maybe instead of many hundreds). Apparently these physicists believe that the superstring theory as an objective description may give us a true picture of the fundamental laws and entities of nature.

However, these physicists take their theories to be understood as an objective account of the constituents of the micro-world. Today, most physicists and philosophers share the same view concerning all scientific theories, though many of the models operate with unobservable objects and properties transcending all possible sense experience. It is commonly thought that the permanency and regularity of observed phenomena can be explained only by postulating the existence of so-called “theoretical entities” or fundamental structures beyond what can be the objects of our immediate experience. Here it is assumed that the business of science demands revealing the laws and hidden structures behind the phenomena, and in doing this, despite our cognitive limitations, science aims to provide us with a true or approximately true description of reality.[11] By and large, science is regarded as able to meet this demand in virtue of well-established methods and practices.

This view of the relation between a scientific theory and its domain of reality seems to have excellent support as long as it gains that support from theories concerning observables like those of classical mechanics. But after the arrival of the theories of relativity and, especially, quantum mechanics in the beginning of the last century—theories departing significantly

[11] I think that the usual interpretation is that science aims at such a description, but has not yet attained it; we have some approximation to the truth, but not the “final” or “ultimate” truth of the matter. The part that I find most problematic about this common characterization of the realist attitude is that it silently moves from asserting that science provides **a** true description of reality to the conclusion that science provides the one-and-only true description. I think one can still call oneself a “realist” and assert that there are a multitude of possible true descriptions—because nature still determines what is true—each appropriate or somehow useful for a particular context.

from the conception of classical physics—an ongoing debate among philosophers has been about whether or not science tells us a true story of the world. It is even reasonable to say that this debate was already going on in nineteenth-century debates over atomism, thermodynamics, and electrodynamics and that the revolutions of the twentieth century just exacerbated this problem. At the end of the nineteenth century, physicists, like Ernst Mach and Wilhelm Ostwald, questioned the reality of what could not be directly experienced such as atoms and fields. Accordingly, hypotheses about invisible entities could not be true, although they might have thought-economic value and some kind of predictive power. However, during the first two decades after the turn of the century almost all physicists came to accept that the dispute was scientifically settled in favor of the atomic hypothesis. Mach's phenomenalist approach to the existence of atoms was therefore no longer a live option among those physicists who, in the 1920s and 1930s, were debating the interpretation of quantum mechanics.

Thus, the classical dispute between Bohr and Einstein was not a debate about whether or not atoms are real. Neither of them doubted the existence of the atomic world. The issue they debated concerned how to describe this world unambiguously. For Einstein "Physics is an attempt conceptually to grasp reality as it is thought independently of its being observed. In this sense one speaks of 'physical reality.'"[12] Bohr, on his side, maintained "Indeed…physics is to be regarded not so much as the study of something given in advance, but rather as the development of methods for ordering and surveying human experience. In this respect our task must be to account for such experience in a manner independent of individual subjective judgment and therefore objective in the sense that it can be unambiguously communicated in the common human language."[13] What distinguished Bohr from Einstein was what each assumed about when a description of atoms can be regarded as meaningful and what the conditions are under which such a description is unambiguous or not.

[12] Einstein (1949), n.3, p. 81.

[13] Bohr (1963), p. 10. The original English text has 'a priori given' instead of 'given in advanced.' However, 'a priori' gives wrong philosophical associations as something existing in the mind prior to and independent of experience. Neither does 'a priori' match Bohr's Danish words 'på forhånd givet,' which means something exists pre-structured or pre-established prior to its experience.

Einstein stood firm with realism as it was built into the foundation of classical physics, believing that statements ascribing properties to atomic objects are meaningful, even when it is empirically impossible for us to determine whether or not such a predication is true. The EPR argument provides an especially straightforward argument in support of this realist notion of truth conditions. In contrast, Bohr believed it is meaningless to ascribe properties to atomic objects if it is, in principle, impossible for us to justify the truth of such a description. The only situations in which the predication of a conjugate variable to an atomic object can be justified are those cases in which we can refer to the value of such an observable in the context of a certain measurement. This is what Bohr meant when he emphasized that the quantum mechanical formalism applies only given certain definite experimental conditions.

Although they followed different lines of reasoning from that of Ernst Mach and Wilhelm Ostwald, some contemporary philosophers argue against reasoning from the purely empirical success of a theory to any metaphysical commitments to imperceptible entities postulated by that theory, and consequently to the idea that empirically successful scientific theories reveal the underlying structures of the world. For instance, their efforts to moor all knowledge to a physicalistic and a non-metaphysical foundation, the logical positivists denied that scientific theories represent any hidden structures of nature, since every claim about theoretical entities could in principle be translated into claims about observation.

After the 1960s the decline of positivism gave an opening for scientific realists, like Karl Popper, to defend the opposite claim in their criticism of the positivists' program. But now the debate moved from a polar to a trilateral format, for at the same time other philosophers like Thomas Kuhn and Paul Feyerabend—equally critical of positivism—argued that the incommensurability of scientific theories disproves the contention that science comes closer and closer to the truth. A generation later, Bas van Fraassen formulated an updated defense of the positivist view freed from its "linguistic straight jacket," which he called "constructive empiricism," and which generated a fierce discussion between the adherents of these two rival realist and antirealist views. Since then, the debate over science's ability to describe reality has seesawed back and forth.

The central philosophical problem in the recent discussion is whether or not there are convincing reasons to believe that all those aspects of a scientific theory that do not refer to anything observable, and thus that are not directly empirically accessible, are a true description of reality. On the one hand, scientific realists argue with a few exceptions that the aim of science is to produce a true or an approximately true description of the world as it is independently of our cognitive powers. Science mainly is concerned with searching for truth. In connection with this assumption, when there are good reasons for accepting of a theory, it is reasonable to believe that it is true. As grounds for their position, scientific realists put forward arguments such as inference to the best explanation, the principle of a common cause, and the success of science. On the other hand, the opposed view of antirealism asserts that the aim of science is to provide us with theories that are merely empirically adequate, that is, theories that only give us an account of what we can observe, just as it claims that nothing other than empirical justification takes part in the acceptance of scientific theories. Usually such an assumption is associated with empiricism. The vindication of antirealist views is commonly based on inferences by historical induction, on arguments for the underdetermination of scientific theories by empirical data, and on arguments for the incommensurability of scientific theories.

Indeed, the current debate is more complex than I indicated above: realists may be further subdivided into those supporting entity realism with the kind of arguments used by Ian Hacking and Nancy Cartwright, those defending theory realism in the manner of Richard Boyd and Stathis Psillos, and those advocating structural realism like James Ladyman and Steven French. In general, entity realists consider themselves as scientific antirealists with respect to theories in contrast to scientific realists, whereas some empiricists do not regard themselves as scientific antirealists. This might, for instance, be Hans Reichenbach and Wesley Salmon. There are really many sides in this debate as characterized by Richard Boyd long ago[14]: empiricist realists, empiricist antirealists, and constructivist antirealists, entity realists and structural realists. Of course, each has its favored argument against the others.

The debate between realists and antirealists has far-reaching influence on our understanding of scientific progress. Realists believe, in general,

[14] Boyd (1985b).

that science comes closer and closer to what is true because of the methods that scientists have at their disposal for making rational choices between alternative theories, while antirealists often believe that truth is relative to a particular conceptual framework, owing to the fact that they have eventually abandoned the earlier idea that there is a neutral observational foundation from which various frameworks can be evaluated. These antirealists also assert that all frameworks are equally acceptable: There is no progress other than greater empirical adequacy. In their most radical forms of social constructivism antirealists become relativists with respect to theory choice, asserting that not objective facts but only historical, social, and cultural circumstances determine which scientific theories are held to be true.

Well, this characterization of antirealism may sound like there is a continuum of antirealisms from Bas van Fraassen to Andrew Pickering, but that is not exactly on target. There is a significant difference—which does not admit of degree—between true empiricists, i.e., those who think theory choice is normally or ideally settled by empirical evidence, and those who think empirical evidence cannot ever settle theory choice and that the real reasons for theory choice are social or economic or what have you. Among the empiricists, some who accept inference to the best explanation, including some pragmatists, think that the best explanation for the agreement between theory choice and evidence is that the theory is at least approximately true. Others, like van Fraassen and other pragmatists, are empiricist antirealists, and they typically find the inference to the best explanation question begging. (Of course, there are or were rationalist realists like Descartes, but I think that species is now extinct, and I guess rationalist antirealists are normally idealists.) But the typology of realists and antirealists, and how they relate to the history of empiricism and rationalism, and to pragmatism, is not particularly obvious or clear-cut. Perhaps some kind of chart comparing the views of Machian phenomenalism, the logical positivists (both realist and non-realist), the Popperians, instrumentalists (or operationalists), constructive empiricists, pragmatists [both realist (Brian Ellis) and not (Richard Rorty)], and constructivists (of a variety of types) is almost impossible to draw. It is really quite a diverse spectrum, which demonstrates that there are not one but many images of science.

At the same time that philosophers of science have been debating realism and antirealism with respect to interpreting the scientific description of nature, there has been a parallel debate among philosophers of language about the same basic issue with respect to what constitutes linguistic meaning and truth. In fact, both, partly separated, debates have their roots in methodologically different approaches to the realist and the antirealist ways of thinking. Realists take their departure from their goal of justifying an ontological description of the world. They then try to elaborate a semantics and epistemology conforming to their metaphysical analysis. In contrast, the antirealists attach much more weight to epistemology and the analysis of the relevant conditions under which we have reliable knowledge, which they subsequently uses to develop both a semantics as well as an ontology. The opposition between these different philosophical agendas reappears in the debate concerning realism and antirealism in the modern semantic clothing in which Michael Dummett has dressed it.

Dummett argues, on the one hand, the antirealist claims that truth is somehow connected with our cognitive abilities, namely what we can legitimately (or justifiably) hold as true depends on the grounds we can cite for believing or accepting it to be true. The ground, which entitles us to hold that something is true, is usually taken to be a body of empirical evidence. So, on the basis of that assumption, the antirealists can claim one of two things: they may consider the truth of a certain belief as something that we are justified in holding either in virtue of our *actual* grounds for holding the belief, whether it is true or not, or in virtue of all possible grounds assuming we have infinite time and perfect epistemic conditions at our disposal. Truth is what is decidable either by direct observation or by a reliable procedure of inquiry.[15] For the realists, on the other hand, this is no requirement. On their view, truth is determined by the relationship between our beliefs and the reality, quite independently of the question of whether or not we possess proper means to decide their actual truth value. Thus, a proposition can easily be true without its content of truth being acknowledged—or justifiably asserted. Later I return to

[15] Unless you already know what is in fact true, one may ask how you could know that any procedure is in fact reliable. The best you could do was to show that for a procedure that made predictions that could be known by direct observation, in the past, predictions made using this procedure were subsequently verified by direct observation, so we are inductively inferring that this procedure will be as predictively successful in the future as it was in the past. Ultimately direct observation is the only epistemic criterion for truth.

Dummett's semantic distinction to see how evolutionary naturalism falls in between realism and antirealism in its treatment of truth and meaning.

1.3 Kant's Metaphysical Dualism

Historically, the metaphysical dispute about our knowledge of reality goes back to the breakthrough of modern science in the Renaissance, and especially Descartes' methodological solipsism, which created a dualism between the knowing subject and the object known. For Descartes positive knowledge of the world is confined to judgments based on clear and distinct ideas grasped by reason. No subject can gain a posteriori knowledge about the world as such, but merely a sensory representation of it that turns out to be quite different from the real world of entities possessing only primary qualities. The immediate object for our sensory cognition is an "internal" representation of what is given in sensory intuition. The culmination of this segregation of the mind and the world is reached with Kant's distinction between reality as it appears to us, organized and structured by the categories of the mind, and reality as it is in itself, transcending any possible kind of understanding. Of course, there is no scientific understanding of things-in-themselves with the faculty of pure reason, but the transcendental ideas have a regulative role. This duality is common to both Kant and the British empiricists. The difference between them is that Hume denies that the content of the mind, in the form of sense impressions, can give us certainty for asserting true belief claims about the nature of the external world, whereas Kant argues that the coordination of the sense impressions by the a priori forms of intuition, together with the subsumption of phenomena under the a priori categories of understanding, gives us genuine knowledge of the external world as it appears for the knowing subject, but not such knowledge of the external world existing independently of this subject.

It is, indeed, debatable what Kant understood by his distinction between sensory appearances, as part of the phenomenal world, and things-in-themselves, as part of the noumenal world. Sometimes he expresses himself as if appearances are not mental entities but material objects as they are known to us, but very often he refers to them as nothing but phenomenal presentations. For instance, when Kant characterizes himself as a metaphysical dualist by being an empirical realist as well as a transcendental idealist, he does so in a way that makes the empirical realist looks very much like a phenomenalist: "Matter is with him [the transcendental idealist] only a species of representations (intuition), which are called external, not as standing in relation to objects *in themselves external*, but because they relate perceptions to the space in which all things are external to each other, while yet the space itself is in us."[16] But if space is in us, then so is matter as pure representation. "In our system… these external things, namely matter, are in all their configurations and alterations nothing but mere appearances, that is, representations in us, of the reality of which we are immediately conscious."[17]

Any interpretation must give Kant benefit of the doubt and make his thoughts as coherent as possible. Taking this principle into account, it seems to me that Kant explicitly denies that phenomena are merely subjective; the phenomenal world is real, not ideal. He clearly thought of himself as an empirical realist. But I would not say that for that reason he thinks that phenomena *are* material objects, for being a material object is itself a concept of the understanding; materiality is subjective, but phenomenal objects are real. This may very well be an inconsistent position—at least history read Kant as an unstable state for philosophy, which quickly decayed into idealism.

Kant's conception of the thing-in-itself (*das Ding an sich*) is equally ambiguous. On the one hand, he seems sometimes to regard the thing-in-itself as a transcendental object independent of our knowledge of it. In that case things-in-themselves should not be regarded as the unknowable 'external' causes of the mental appearances, since this would be an application of the concept of causation outside the sensuous range of understanding. On the other hand, time and again he speaks about things-in-themselves exactly in this fashion. "The understanding in limiting sensibility…thinks for itself an object in itself, but only as a transcendental object which is *the cause of appearance* and therefore not itself appearance, and which can be thought neither as quantity, nor as reality nor as substance, etc."[18]

[16] Kant ([1787]2007), A 370.
[17] Ibid., A 371–A 372.
[18] Ibid., A 288/B 344.

Afterwards one cannot see how Kant can say something like this without believing that he needed the thing-in-itself as a causal counterpart to the mental appearance. Kant addressed Hume's metaphysical skepticism by claiming that the phenomenal world, in contrast to the world-in-itself, is not ontologically different from our mental representations formed by applying the categories of the mind to the sense impressions. Nevertheless, Kant argued for the existence of things-in-themselves based on metaphysical reasoning: partly, as Descartes had done before him, by assuming that the cause of our mental presentation has its 'external' grounds in something that is not directly sensed and partly as a limiting notion of a thing that is not an object of sensuous intuition.

Perhaps it is not fair to say Kant was a phenomenalist, which would make him a skeptic about the *reality* of an external world. 'Matter' is also 'in us' because it is, in effect, a category of the understanding. The key notion here is not 'matter' but 'externality.' The thing-in-itself, of which we can form no representation, is by definition the *external* thing that grounds the phenomenal object as not merely subjective. I think the best we can say for Kant is that he did not consider the "grounding" relationship between appearances and things-in-themselves as akin to what he called "causality," because for him causality was in effect Newtonian, mechanical causality between presentations in space and time; thus, when he speaks of the noumenon as the grounds of the phenomenon he has in mind something quite different from what he called "causality." However, Kant's project is quite different from most people today who would accept the point of his Copernican revolution. To show how synthetic a priori knowledge is possible Kant has to prove the universality and necessity of his concepts; no Darwinian-oriented philosophers would undertake such a task. The concepts are neither universal nor necessary but the product of natural selection, which exhibits neither of those characteristics. So Kant's strictures on the use of causality (which he violates in the passage just quoted) need not apply to today's naturalist/pragmatist outlook. However, according to this outlook, Kant is not justified in believing that something of which he can form no representation exists.

Contrasting with his characterization of the transcendental idealist, Kant placed the transcendental realist, who also is a dualist in virtue of being an empirical idealist: "After wrongly supposing that objects of the senses, if they are to be external, must have an existence by themselves, and independently of the sense, he finds that, judged from this point of view, all our sensuous representations are inadequate to establish their reality."[19] But because both of them are dualists, then neither the transcendental realist nor the transcendental idealist is able to establish the one-world view in which our sensory representation of the world is also the external world as we possibly can represent it. To overcome dualism, and the skepticism following from it, we need a philosophical argument according to which reality at least includes the phenomenal world as we experience it.

Modern philosophers have generally proposed two opposed strategies in order to satisfy the Cartesian quest for finding a reliable procedure of inquiry into the external world. The first suggests that science, in particular through relativity theory and quantum mechanics, has overcome the Kantian distinction. According to this strategy, these sciences show us that part of reality lies behind what can be seen with the naked eye. Especially the use of mathematics in science yields a representation of how things really are independently of our experience of them. Physics need not rely on any sensuous representation in order to characterize the external world. The world of modern physics is the real world as it is in itself. The other strategy claims that reality is by necessity restricted to what in principle can be perceived by human sense organs, a view that in its ultimate and most extreme form restricts reality to our subjective impressions (subjective idealism) or the immediately given phenomenon (phenomenalism). However, neither strategy makes much sense from a Darwinian perspective because both mathematics and sense impressions are results of human evolution.

1.4 Evolutionary Epistemology

Over millions of years, biological selection adapted hominids' cognitive powers to meet the sensory input from the external world. Before the evolutionary tracks of hominids and the other primates diverged, most of

[19] Ibid., A 369.

the sensory mechanisms, including those that detect colors, had already evolved. Other mammals have little or no color vision. Most of them, like dogs, for instance, are dichromats, having only two types of cone cells in their retina, in contrast to humans, who are trichromats. This enables the dog to distinguish blue from yellow, but not red from green. However, this does not imply that a red berry is not really red or that red berries are not sweet and juicy.[20] For human beings colors are empirically real properties. They make a difference to us in how things really are in our sensory experience and in the way we perceive, act, and think about them. Red berries can be seen on a background of green leaves and signal readiness to be picked and swallowed. But according to contemporary scientific theory there is more to the story than this. Colors appear as they do because of the way we are adapted to present to our consciousness information about the surfaces of the objects in our environment caused by the interaction between our optical sense organs and electromagnetic radiation. Investigation has shown that colors do not have the nature that naïve realism attributes to them. Naïve realism treats colors as having a simple directly experienced nature; we simply see that something is red; we do not get it from something else. However, scientific inquiry reveals that colors are the causal result of how things' surface structure appears to us in the presence of light and the properly functioning neurons in the human brain.[21] Colors are epistemologically simple, but ontologically complex. Some might take this to indicate that the sciences deal with Kant's colorless things-in-themselves after all, but such a conclusion does not follow.

For science to establish that colors can be associated with how surfaces reflect light waves, and that color qualia may be associated with our sensory capacity to transform electromagnetic radiation into perceptual information, one cannot strip colors from the scientific image of nature.

---

[20] One might ask which red is real, the physicist's red as the ability to reflect light of a certain frequency or the empirical red delivered by my senses? However, there is no physicist's red but only a physicist's and a neuroscientist's explanation of why red is empirically real.

[21] This train of cause and effect takes us only as far as the excitation of neurons in the brain. It is another step from that neurophysiological event to the red qualia that enters into my consciousness; there is nothing that is "red" in my brain, but there is a very real empirical red experienced by my consciousness (a.k.a., my "mind"). This is the big issue, though not entirely unexplainable in my opinion (Faye, 2013).

These qualities are just as real as anything else scientists work with. In the right season cherries have green leaves and red berries not only for the scientifically unschooled, but also for the horticulturalist and the evolutionary botanist. It would be absurd to suggest that colors, sounds, tastes, and odors do not exist in spite of the fact that the human species is adapted to experience these qualities. This much we are told by biology. It is only when science begins to ask questions about how perception of colors is physically caused that color is no longer glued to the surface. The physical sciences tell us that perceptual qualities are much more complexly created than naïve realism imagined. Even so, the physical sciences have to carry out their investigation within the scope of our naturally evolved sensory capabilities including the ability to see red berries in the proper light.

Humans and chimpanzees have between 98 and 99 percent of their genes in common. This means we share with chimpanzees many physical similarities as well as cognitive similarities. The shared cognitive mechanisms include (1) non-linguistic concept formation, i.e., thinking about kinds as well as numbers; (2) simple addition and subtraction; (3) imagination; (4) inductive reasoning; (5) deductive reasoning; (6) causal understanding; (7) error finding; (8) spatial-temporal awareness, (9) a realist instinct, i.e., the capacity of distinguishing between bodily internal from bodily external information, and perhaps (10) a theory of (other) minds, i.e., knowing both one's own intentions and the intentions of other members of one's species.[22] These capacities have evolved

[22] See, for instance, Boysen (1993); Boyson & Hallberg (2000); Vlamings, Uher & Call (2006); and Matzuzawa (2007). Animals and birds may have intentions without being consciously aware of these intentional states. Some researchers, like Hare, Call &Tomasello (2001), believe that chimpanzees possess a theory of mind as an intervening variable such as being aware of their own intentional states as well as imputing intentions and knowledge to others; other researchers like Povinelli and Vork (2004) claim that all data can be explained as behavioral abstractions because of separate associations between singular types of perception and singular types of behavior. See Shettleworth (2010) pp. 441 ff. for a thorough review of the research data in this area. As she correctly remarks, "In the classification of intentional states, theory of mind implies second-order intentionality" (p. 441). However, I think from an evolutionary point of view it is much more behaviorally efficient to be aware of one's thoughts and wishes as an intervening variable between different types of thoughts and different types of behavior rather than to behave on association of singular type of thoughts with singular types of actions. It is, of course, an empirical question whether or not an animal possess self-awareness and other-awareness, but I would be surprised if at least the great apes (and perhaps birds like ravens) do not possess some limited awareness of their own mental states as well as others.'

adaptively as cognitive mechanisms for handling sensory information.[23] What we and the chimps do not have in common is an advanced capacity of reflective thinking, the use of language for communication, oral as well as written, and an advanced capacity of mathematical reasoning. Reflective thinking and language gave human beings not only the ability to contemplate their own cognitive practices, but also the ability to reify or hypostasize their own thoughts. Some researchers believe that many components of human linguistic ability can be found in other species as part of their natural communication system, yet in the evolution of human language these components have become uniquely co-adapted for linguistic communication.[24]

[23] There are two possible interpretations to what I am saying: (1) human and chimp cognitive powers are 'homologous,' i.e., derived from a common ancestor, and this explains their similarity, or (2) the two are 'analogous' or 'convergent' where evolution selected the two forms separately because they both proved positive adaptations for tasks the organism faced, i.e., the similarity occurs not because of common ancestry but because of common needs and environments. I think it is combination of the two, partly because chimp cognitive powers will also have evolved after the separation from their common ancestry with humans, but this is all together an empirical question.

[24] Chomsky (2004) and (2005) has suggested that the evolution of language was a very discontinuous process taking place as a single mutation around 100,000 years ago. For him, language ability consists of being able to construct and understand recursive data. Hence, Hauser, Chomsky & Fitch (2002) find that this ability defines the faculty of language in the narrow sense (FLN), whereas all perceptual, motor, and cognitive abilities that contribute to language but are shared with other species belong to the faculty of language in the broad sense (FLB). It has indeed been shown that the protein FOXP2 affects language learning as malformation of the FOXP2 gene causes loss of language skills, but the protein is to be found in animals as well as songbirds. The protein in humans is distinct from those observed in chimpanzees by the substitution of two amino acids, in mice three, and in songbirds seven amino acids. Others such as Pinker and Jackendoff (2005) point out that this disposition evolved more gradually on the basis of many mutations and that the precursors of language in general are not to be sought among non-human primates. Here it is worth noting that ever since chimpanzees and bonobos shared a common ancestor with humans about 5–7 million years ago, natural selection has been effective among their as well as our ancestors. So it is not at all certain that their manner of communication and our language ability are genetically linked particularly closely. The Danish linguist Ib Ulbæk (1998) thinks that language has not evolved from earlier primates' ability to communicate, but from their cognitive abilities, i.e., as a non-linguistic understanding of their surroundings in the form of remembrance, recognition, and performance. In my view, Ulbæk correctly suggests that cognitive development from early primates to *Homo sapiens* has proceeded fairly gradually, while language ability has evolved in spurts. This hypothesis fits well with the notion that concept formation phylogenetically precedes language formation. In order for language to develop, a number of anatomical and physiological traits had to be in place. Upright walking not only freed our human ancestors' arms and changed the shoulder blades, so we can throw overhand, but it also opened the pharynx and oral cavity. Today, it is estimated that about 100 different muscles are included in speech from the air-

So the result of the cognitive adaptation of human perception was later embodied in the basic rules and categories of human languages. However, in order to be of any use for interaction with the environment, perceptual information had to be processed and compared with earlier stored information. The ability to compare sensory information had a higher survival value if it was not always carried out automatically but sometimes happened deliberately by conscious reflection. As a result perceptual concepts became blended with concepts created by our ancestor's imagination. The development of a language of perception supplemented with a language of reflection eventually caused us to grasp the world in metaphysical terms of space, time, thing, event, property, and causality and to structure our description of the world in the way in which these terms were interrelated by our basic conceptual framework. Today we use our natural language to express our particular beliefs about everyday states of affairs. Languages as such are neither true nor false, but they supply us with the power of making true or false assertions about the entities posited by the senses as well as about what we directly sense. In order to make such claims, we need some linguistic exit and entry rules (the arrows of reference), which identify the spoken-about entities, and those rules are therefore analytical although originally empirically attained.

As a Darwinian I claim that biology reveals the grounds for an ontological understanding of our everyday knowledge. The existence of the reality we experience proves itself conditionally by taking Darwin seriously, as Michael Ruse urges us to do.[25] Common-sense realism is justified by its immense survival value. If we did not trust our senses and the information about an external world that they reveal, we would have been gone long ago. Our sensory capacities, and the "trust" we place in them, co-evolved; one could not have evolved without the other. While natural selection does not determine that an organism will evolve in a certain way (because variation is "random"), it does bar certain possibilities from ever evolving. Evolution by natural selection, including the evolution of the cognitive mechanisms of the human species, implies that the capabilities of our conceptual and

ways to the lips. In addition, many social competences had to be in place as well before the ability of language could evolve.

[25] Ruse ([1986]1998). The title of this book is *Taking Darwin Seriously*.

linguistic resources were constantly adapted to material conditions outside our body and mind. Consequently, the distinction between an external world existing independently of any human being and the thoughts and beliefs that human beings hold about this world naturally evolved. I hold that this distinction between what refers and what is referred to is a precondition for a theory of meaning for any natural language. Knowing the meaning of a statement in any such language is equivalent to being able to form what the community, typically, believes to be a true sentence about reality, and knowing reality (as well as knowing one's thoughts and beliefs) seems to be equivalent to the ability to state what the community takes to be the appropriate true statements. Yet, it is open for debate how far meaningful statements about nature can transcend possible experience and still truly express entirely objective states of affairs.

Indeed, we need an ontological distinction between a knowing subject and the object known simply because our thoughts and beliefs are logically distinct from what they are presumably about. Although no subject may be thinking about or perceiving an object, this does not imply that it does not exist. Nor does thinking of it, or even allegedly perceiving it, imply that it does exist. The dichotomy between the external world and the species-specific apprehension of it can be overcome by virtue of an ontological commitment to the externality of the cause of what epistemologically may appear to be its internal effect. This ontological commitment requires that a Darwinian epistemology replace the Cartesian epistemology. We are biologically adapted to makes a distinction between entities that are regarded as external and as internal. For, as a matter of fact, a normal (i.e., non-hallucinatory) sensuous *presentation* is comprehended under the category of material objects (and not sensory appearances). This category has evolved in pre-scientific cognitive practice and eventually is expressed in terms of our natural language. Natural language has evolved as a means of communication that enables us to express our ideas, thoughts, and beliefs at the same time as such expressions may also designate an external world. It has the power to express our immediate experiences, while at the same time it signifies what this experience is about. So the use of a natural language commits us ontologically to an external world, as the world is described in terms of what we experience it to be.

So the traditional epistemology, which we can trace back to Descartes, distinguishes between subjective and objective knowledge by claiming that the subjective and the objective are not the same, nor do they have the same properties. Subjective knowledge is about the mental apprehension of an object that is *caused* by the non-apprehended or 'external' object. In contrast, objective knowledge is apprehension-independent. It reflects the object as it really is by describing it in terms of definite self-determined properties. But Darwinian epistemologists deny all this tradition. To them, the distinction between objective and subjective knowledge is disastrous. Instead, they maintain that natural selection has empowered us with the ability to separate an external world from an internal world, but it is still such that the external world is grasped in terms of species-dependent categories just as much as the internal world. According to them, we cannot have knowledge of a species-independent world as it exists absolutely in and by itself. Whatever knowledge we have of an external world must depend on who we are and be based on how we are adapted to experience it. Regardless of whether we describe the external or the internal world, we may give a mind-independent description of it, but that description, because it is *we* who give it, must necessarily be species-dependent.

The distinction between an objective, mind-independent domain of things-in-themselves, on the one hand, and a subjective, mental, world of sensory appearances, on the other hand, so saturates the Cartesian-Kantian tradition in epistemology that it can be extremely hard to formulate any alternative. (Actually the ancients already had an alternative, which has no reference to the objective-subjective dichotomy; it is Aristotelian science.) The evolutionary naturalist is a common-sense realist distinguishing the external from the internal world and by claiming both that the world we perceive to be separable from our direct experience of it is external and that the world we perceive to be inseparable from our direct experience of it is internal. We may say then that external things experienced to be separable from consciousness have an "objective" or "mind-independent" existence. If this is the full meaning of these terms, it seems quite unproblematic. However, the evolutionary naturalist also maintains that all knowledge of the external world as well as of the internal world is species-dependent. If the terms "objective" and "mind-independent" are used to refer, inclusively or exclusively, to objects beyond our species-dependent power to observe, then the evolutionary naturalist would say that only things-in-themselves can have that status. But we cannot have knowledge of what we do not have the power to somehow observe, since we are adapted to have only that information we come to possess as a causal effect of the interaction between the electromagnetic field and certain neurons in our sense organs. So, given the species-dependent nature of all human knowledge, the notion of a thing-in-itself apart from how any organism cognizes it does not really make sense from a Darwinian point of view. The Darwinian insists that if 'objectivity' is taken to be equivalent to 'mind-independent,' and the latter is considered to be equivalent to 'species-independent,' then no human being can possess that kind of 'objective' knowledge. However, 'objectivity' and 'mind-independency' are sometimes associated with 'intersubjectivity,' in which case we mean only that this knowledge is independent of any particular individual mind. This raises no concern for the evolutionary naturalist since all knowledge that is intersubjectively established is species dependent. In this cognitive sense, both the external as well as the internal state of affairs can be said to 'exist objectively' in the world.

Nevertheless, science makes extensive use of abstract concepts, which seem to have no direct connection to our experience of an external world. Do not at least some of these concepts stand for a representation of an objective reality beyond an experiential presentation of the external world? I would take some of the abstract concepts that are (normally) believed to represent an external world to be substantive concepts representing entities, their properties, or their behavior, e.g., atoms, electron spin, or pair annihilation. But truth is also an abstract concept. 'Truth' is a "second order" concept in the sense that it is about the relation between these substantive concepts and the things they may or may not represent. I shall argue that our evolutionary heritage puts a restriction on the scope of a realist interpretation of science. I hold that the concept of truth contains a certain epistemic element because the concept itself is an abstraction from our innate ability to distinguish functionally between successful and unsuccessful beliefs: Thus, what makes a belief true or not depends on our ability to establish the existence of a truth relation between that belief and what makes it true. I contend that the truth relation is external rather than internal: external states of affairs make our beliefs true or not, only because they can in principle be demonstrated, only because such states of affairs can be compared with other states of affairs. In all other cases, our beliefs are about abstract mental constructions. How this works out in detail will be explained later.

Thus, in contrast to the classical debate between realists and instrumentalists, I claim that the contemporary metaphysical discussion of science should not target the question of whether or not so-called "unobservable" entities are real. My suggestion is that from a Darwinian point of view there is no difference between observing everyday things, like human beings, horses, and cars, and instrumentally observing scientific objects like fields, electrons, atoms, or viruses. Clearly, we are not adapted to see what is invisible to us. Here technology compensates for our lack of the appropriate sensory channels. Of course our belief in such invisible entities is highly dependent on our state of knowledge at any given time, but the observation of visible entities is just as much knowledge-dependent as is the observation of invisible entities. The proper metaphysical debate should be about how claims about invisible objects become true given the cognitive mechanics that we have selectively acquired to cope with our environment. The unfortunate distinction between visible and invisible entities is founded on a dualist view of representational knowledge. A common way of looking at the distinction is to think of the visible objects as the objects of sensuous representations and to think of non-perceptual objects as belonging to the noumenal world-in-itself. Nonetheless, I claim that this is an unfortunate way of looking at the problem. Whenever we make any claim about either visible or invisible objects, we cannot separate the question of what endows these presumably different statements with a truth value from the question of which epistemic constraints these different assertions may be subjected to.

As we have noted, from Descartes forward the epistemological struggle has been about how we could justify any idea or sensation as a faithful *representation* of the world or how we could justify which ideas or sensations are faithful and which are not. For a long time, it was agreed that we need some concept of representation in the analysis of the relationship between our perceptions and the world. Ideas and sensations are different from what causes them; physical objects are logically distinct from the mental states that point to them. Ideas and sensations took the place of the real things: they became an internal substitute for external things.

God was, perhaps, the only being who may not have such a relationship to the world; in this case the world was regarded as nothing but a pure manifestation of his thoughts. However, nobody can compare his or her mental representation with the world-in-itself; one can only compare one idea with another idea, one sensation with another sensation. So how can we know whether or to what extent our ideas and sensations resemble the world-in-itself? Even if we hold that they do not have to be similar to what is represented, it is still unclear in which sense the so-called "appearance" is epistemically separable from the object of experience. Within the scheme of mental representation itself, we cannot distinguish between form and content. Hence, for a Darwinian, it becomes meaningless to make an epistemic distinction between what cannot in practice be separated.

The principles of scientific reasoning or methodology have their origin and only justification in this evolutionary process of selection. The basic cognitive mechanisms, by which human beings process external information, were selected according to their functional efficiency.[26] Later, when evolutionary selection empowered human beings with a capacity of reflection, we were able to investigate consciously our own experiential and cognitive practices, and this gave rise to these principles of scientific reasoning. Similarly, scientific theories as expressions of conceptual frameworks provide us with an exact language with which we can communicate scientific information in a very efficient way. As such theories are neither contingently true nor contingently false, they can, however, be used to make true or false statements about posited entities and the causal laws governing them. Thus, scientific theory offers a language that makes the description of such entities intelligible. Here the fundamental laws provide us with the same exit and entry rules of the theory, since these laws are nothing but rule-giving statements that at the same time are given as analytically true, but can be known only a posteriori.[27]

The focus on a Darwinian epistemology explains why the sciences vindicate their discoveries and assumptions as they do, as well as why

[26] Natural selection has to work with the materials that are available at that stage of evolution. What may evolve may be functionally the most efficient *given what nature has to work* with. If natural selection had been "given" different materials to work with, different—and perhaps functionally more efficient—capabilities or organs may have evolved.

[27] See Faye (2014), p. 103 ff.

scientists believe they possess knowledge of nature and are ontologically committed to the world as described by natural science. Both ordinary experience and science belie philosophical doubts about realism because our senses and powers of reasoning are selectively adapted to handling information about the external world such that this information becomes useful for actions and reactions. A Darwinian epistemology makes us common-sense realists and explains why this is so. If we subscribe to a Cartesian epistemology, and attempt to explain our beliefs in an external world, we face an explanatory gap between sensory appearance and a reality that is assumed to lie behind the appearance. A Cartesian analysis requires us to try to prove the reality of a world that our basic instincts assure us exists. But for a Cartesian the true reality is the one given by reason and not by the senses. The consequence is metaphysical skepticism concerning the ability of the senses or empirical inquiry to deliver true and warranted beliefs—given the failure of Descartes' own strategy via the veracity of an infinitely Perfect Creator. An evolutionary epistemology is the viable alternative that explains our capacity of knowing the world in terms of natural selection. But in no way is it straightforward how evolutionary epistemology can guide us in our search for an evolutionary ontology that also underpins scientific research. Because, as Michael Ruse explains to us: "We are animals using our evolutionarily acquired powers to delve into questions for which such powers were certainly not intended."[28] The question is not whether scientific beliefs represent external states of affairs, observing electrons are not part of these powers, but how far these beliefs can be said to represent the world as it truly is. Does the nature of scientific beliefs reach no further than what we are justified in saying stems from the representative capacity that selection has bestowed on us? Or to put it in more general terms: Which beliefs do we impose upon experience and which do we draw from it? This question awaits a decisive answer.

Perhaps Cartesianism leads to metaphysical skepticism, but if that is so, then it seems Darwinism does too, since it denies that we can have knowledge of a species-independent world-in-itself. Elsewhere I have distinguished the distinction among three metaphysical positions: realism,

[28] Ruse ([1986]1998), p. 188.

antirealism, and agnosticism.[29] I defined these positions with respect to their contentions about things-in-themselves and the kind of knowledge we may have of them. If we follow Ruse's interpretation of Kant's things-in-themselves, a Darwinian philosopher could not accept this kind of entity. Ruse believes that the Darwinian epistemology implies metaphysical skepticism.[30] In the present context, this skepticism can be associated with metaphysical antirealism. I think this assertion is generally correct. As we have noticed, Kant's concept of things-in-themselves is notoriously ambiguous and difficult to reconcile with his general strictures on the application of the pure concepts of the understanding. According to one reasonable interpretation, things-in-themselves are objective but *abstract* entities, which do not exist in space and time—hence the noumenal nature of their existence. The only entities a Darwinian will accept as real are those that exist in space and time, and these external entities are nothing but things-as-they-can-appear for us. Things-as-they-appear to us have both secondary qualities as well as primary, but the external things, according to classical materialism, have only primary qualities, and the secondary qualities are added by us.

In another interpretation things-in-themselves belong to an objective and species-independent world consisting of nothing but *concrete* entities since these entities are what cause our sensory appearances of them. These objects could be radiative fields or centers of force as well as atomic and subatomic particles. Therefore, their characteristics may be different from what they immediately appear to be. So based on our experience we

[29] Faye (2002), pp. 209–114. Here I associated metaphysical realism with the claim that things-in-themselves exist, that statements about these things-in-themselves are either true or false, and that we can obtain (a priori) knowledge of their truth value, whereas the metaphysical antirealism was associated with a denial of the ontological, the semantic, and the epistemological component of metaphysical realism. However, metaphysical agnosticism holds that some ontological statements about things-in-themselves are true or false, and we may have knowledge of these truth values. I further separated metaphysical agnosticism into a weak and a strong position. Of those two I defended a strong agnostic attitude, which I defined as the view that "denies that we can have true beliefs about things-in-themselves unsupported by any empirical underpinnings." Today I would state my opinion differently. We can observe invisible objects only as long as we can make them observable to us. Hence, we do not know them as things-in-themselves but always as things-as-they-are-experienced by us. So I agree with the metaphysical antirealist with respect to things-in-themselves. There are no such things. But I am a realist with respect to truth. There are external truth-makers beyond our sensory experience.

[30] Ruse ([1986]1998), pp. 192–196.

may sometimes have good grounds for believing in such entities that are distinct from their visual impressions. Things-in-themselves would then count as all those objective things that cannot be seen by the naked eye but can be observed or, as we often say in science, 'detected,' by instruments. If the Darwinian philosopher goes with this interpretation, she or he would not be a metaphysical skeptic but rather a metaphysical agnostic.

However, no interpretation can overcome a fundamental ambiguity. The external world contains visible as well as invisible things; this is the world we can either experience or observe or detect, and represent, and this is the world of which we can have knowledge. An objective world, in the sense of a world described in species-independent terms as it is in itself, is not something that we are adapted to understand, and therefore is not something we can have knowledge about, and therefore not something whose existence we have reason to postulate. So there is nothing to be agnostic about. Nothing in the account of evolutionary epistemology gives any reason for the metaphysician to favor agnosticism over antirealism.

2

Evolution and Human Cognition

As we all know, Darwin explained the evolution of biological organisms in a way that places human beings as one species alongside other animals. This is an explanation that most of us accept. The consequence of the explanation is that our present cognitive and physiological aptitudes were established in our long-gone ancestors and have been inherent in their successors ever since. The evolution made up our cognitive and physiological possibilities, but evolution also set the natural limits for how far these innate aptitudes can be used. All cognitive and physiological characteristics are due to biological selection and adaptation to a physical environment on which individual organisms in general have little influence.[1] What is strange, however, is that many philosophers, including some well-known contemporaries, ignore this Darwinian legacy. They continue to do philosophy as if our sensory of experience and capacity of cognition

[1] Perhaps individual animals do not have much impact on the environment, but certainly there must be many cases where the behavior of large animal populations has had deep impacts on the environment. In any event, my general impression is that the view that sees evolution as happening within a relatively stable environment is sort of giving way to an outlook that sees organisms as simply more of the myriad parts of the environment itself as a whole complex system, all of which are in dialectical interaction with the others—both organic and inorganic—in myriad ways, which we are barely beginning to understand. Ultimately the whole question of a border between the organism and its environment seems somewhat arbitrary.

J. Faye, *Experience and Beyond*,
DOI 10.1007/978-3-319-31077-0_2

were not the product of organic evolution. Instead these philosophers and philosophers of science occupy themselves with a conceptual analysis of propositions neglecting the fact that these propositions have a human origin.

The Darwinian explanation of human cognition, however, understands cognitive dispositions as primarily a product of biological evolution, which among humans eventually originates in a (co)-evolution of social and cultural behavior with human organismic capacities. The selective development of human cognition has conditioned conceptualization, language, and the very nature of human subjectivity, not the other way around. Neither human consciousness nor the logic of language can be considered as a precondition for human cognition. Hence the Darwinian legacy should warn us against that philosophical hubris of which we can easily become a victim, if we naively believe that science gives us some kind of access to the world-in-itself and that there is a one-to-one correspondence between the scientific representation of nature and nature herself.

In this chapter I shall defend the view that human cognitive capacities should be understood as a result of biological evolution and therefore that the scope of scientific understanding is limited by the sensory conditions to which our conception of the world has become adapted. If this is true, it puts severe cognitive limits on how far beyond our sensory experiences a realist interpretation of science is justified. In the end I shall argue that much of what science and philosophy have to say about the world is due to abstraction and idealization and exists only as a construction of the reflective mind.

2.1 The Darwinian Legacy

When Aristotle characterized humans as rational animals he was closer to the truth than many of his medieval successors, not so much because of our capacity of rational deliberation, which some of us unfortunately do not share, but because he considered us as a kind of animal. It was not until Charles Darwin that we got an explanation of why we share anatomical and physiological features together with many animals. According to Darwin, the human species has evolved over time from ancestors who were non-reflective animals. Hence: once an animal always an animal. This evolution can be described by three conditions that not only developed our bodies but also our consciousness and cognitive capacities. It explains human cognitive abilities as a result of a long process of variation, selection, and retention. By making the mechanism of knowledge acquisition a result of evolution, biology becomes a critical source for understanding human cognition. The branch of naturalistic epistemology that focuses on biology was dubbed "evolutionary epistemology" by Donald Campbell.[2]

Evolutionary epistemology is a form of naturalized epistemology in which the focus is on understanding the cognitive capacities of biological organisms in terms of how these capacities developed under the processes of natural selection. Due to this development, well-adapted organisms have been able to cope with their surroundings, as long as these are stable. According to some evolutionary epistemologists, this implies that the universal evolutionary mechanism, by which organisms are created, is the same mechanism working within the evolution of cognition and within the products of cognition such as language, culture, and society. This universal mechanism is adaptation by natural selection. An evolutionary approach to epistemology is expected to answer all epistemological questions in virtue of biological organisms' cognitive adaptation over time. Defenders of adaptation by natural selection, like Donald Campbell, strongly believe that all biological cognitive capabilities are stored immanently in social and cultural development. In his work Campbell has tried to understand all forms of knowledge from a neo-Darwinian perspective, where selection is the key factor in cognitive evolution because it explains how our cognitive representations are first and foremost adapted to the physical environments, but also eventually to the social and cultural surroundings. Natural selection works such that among variations in a set of similar organisms those that fit best into the actual environment have the greatest chance of survival. An important thing to note is that the adaptation of organisms, according to this model, does not evolve because of instructions from the environment. The model is based on two conditions: (1) Knowledge is regarded as the result of a selection process that generates and maintains the reliability of the senses and the cognitive

[2] See Campbell (1974).

mechanisms that process the information received through the senses. (2) Development of methods of reasoning such as induction, "trial and error" learning, deduction, and other methods, as well as development of scientific theories, is also controlled by selection processes. In other words, evolutionary epistemology holds that the possibility conditions of cognition exist in biological development. Thus this epistemology claims to be able to explain the results of this evolution.

However, in my opinion, Campbell and other adaptationists seem to overlook the fact that the cognitive process of advanced learning acts not only in virtue of a passive adaptation. The important fact, which Campbell's approach ignores, is that the organism must deliberately act on information received, as this is the only way that the organism "learns" whether its representation is successfully adapted to what is represented. The implication is that the individual organism cannot *learn* from the environment unless natural selection has given some of its predecessors an ability to understand sensory information as information about its environment. Induction, for instance, is not capable of working as a leaning mechanism if it merely consists of *trial and error* attempts. An organism learns next to nothing from an accidental behavior that fails by acting in yet another accidental but unsuccessful way. It is time consuming, uneconomical, and may be dangerous, too. A miscarried attempt and one may not get a second chance.

As a form of naturalism, evolutionary epistemology is fully committed to empiricism, but at the same time it rejects the idea that perception is experientially immediate (direct realism). The form of evolutionary epistemology, which I advocate, displays a selection theoretical model of how any given perceptual belief and its object can fit together. According to Campbell, language is neither the foundation nor a prerequisite for cognition. It is a mistake to understand cognition as having a propositional form or as being limited to humans. Rather, all organisms can show behavior that reflects their cognition of their environment. At first what matters is not the issue of linguistic sense, but how we humans can share common linguistic reference and convey linguistically formulated assumptions. In the evolution of human linguistic capabilities the problem of reference comes temporally before the problem of meaning. The ability of reifying our experience precedes the development of a language. In my view this amounts to saying that concept formation (at least up to a certain level) occurs earlier than language formation, i.e., sensory-based concepts are pre-linguistic necessary conditions for language formation.[3]

An appeal to the biological theory of natural selection cannot be used directly to justify scientific (and other more advanced) conjectures, because in this case concept formation goes with language training. This requires an epistemological relevance theory of language learning. Such a theory must start with the assumption that the ability to speak any language is a capacity innate to humans, while the language any individual actually speaks is learned empirically and changes from one culture to another. A number of empirical studies suggest that, for example, our basic color separations fit into all languages, a phenomenon that can be explained only by the fact that our ability to identify colors was established in the evolutionary history of our lineage long before our capacity for language.[4] Empirical studies show color categorization exists in monkeys, birds, and some fish. Noam Chomsky has also argued that all human languages are characterized syntactically by a basic common structure that is innate. Similarly, writers such as George Lakoff and Mark Johnson have pointed out that there are certain fundamental semantic figures of speech across all cultures, unlike the variety of natural languages that displays a plethora of differences in meaning and structures. Not surprisingly, human languages are a product of culture as much as nature. Many of our reflexively acquired assumptions are more closely related to linguistic meaning than to linguistic reference.

Overall, Campbell saw selection in one form or another as the active mechanism of human cognitive development. Therefore he appealed to social selection rather than natural selection in order to solve the adaptation problem in a cultural context. The analogous problem for biological evolution is for example the biological form of organisms. The theory of natural selection denies that organisms' cognitive skills can develop through external "instructions" to the organism. It is only the environment that selects between different phenotypes. The capacity of sight developed from a natu-

[3] See also Ulbæk (1998) who argues that our language has not evolved from prehistoric primates' communication system but from their non-linguistic understanding of their environment.

[4] Cf. Franklin *et al.* (2005); and Franklin *et al.* (2008).

ral variation among phenotypes (generic proliferation and mutations in the genotype) in a number of organisms among which the environment then “selected” some favorable variants that exhibited a phenotype with the capability to detect and focus light and so had a better chance to survive in the environment until they could reproduce. In this way, eyes gradually evolved because their function of providing useful information about the external world enhanced the individual organisms’ ability to survive and reproduce.

Of course, the function of the eye is not always optimal; thus it can easily fool us. Therefore some beliefs acquired perceptually may be suspect. We do not trust these perceptions, because we trust in others—those perceptions that do not betray us are just not erroneous and therefore can be relied on. Perceptual beliefs hang together and therefore are reliable because of this connection. No one of them is individually infallible or basic but all form a system whose individual elements may eventually be revised. Yet we can have confidence in most of our perceptual assumptions, because their relation to other perceptual assumptions establishes their reliability. Therefore, we can have confidence in the majority of such beliefs, while we can also revise a few of them. Yet, if certain perceptions deceive us it is because they lead to a misguided behavior.

But is natural selection a reliable mechanism for explaining how we can have scientific knowledge, i.e., theoretical knowledge that goes beyond what we can see with the naked eye? We must first consider how our beliefs may adapt to what we cannot see: how can it be that in some cases where we cannot perceptually observe the world, it still seems to be possible to develop knowledge that fits nicely with the worlds as revealed in experimental science? The question can be answered only if we can reason from an individual biologically based epistemology to a biologically based social epistemology, which tells us how theoretical knowledge about invisible entities becomes part of our scientific practice. In this situation the environment no longer selects individual organisms, but where empirical facts discovered by scientists act as a selective force on the body of collectively accepted beliefs.

One proposal for understanding this kind of cognitive selection is provided by Karl Popper’s falsificationist methodology. He claimed that empirical observations can never verify or even provide inductive confirmation for any scientific (or other general) assumptions, but that scientific theories and hypotheses can be falsified by observations. In this way observation offers a mechanism for selecting that theory that is most likely to be true. In a manner reminiscent of the Darwinian mechanism of selection, our hypotheses about the world 'adapt' to the world through falsification. As long as they are not falsified, one has every reason to believe them. The hypotheses that are most fit are those that survive even the most demanding tests to which science can expose them. Therefore to show their value scientific conjectures must be constantly exposed to new serious tests deliberately designed to falsify them. Popper's example of evolutionary epistemology is grounded in his rejection of induction as a logical method to identify true universal assumptions. Such universal presumptions arise, rather, as a result of the particular researchers' bold guesses. Ideally the individual guesses of many different scientists build up a pool of variations among the guesswork of the research community, just as new mutations create a series of variations in a population of genetic material. Falsification ensures the selection of the most viable assumptions very similar to the environmental selection of the most viable organisms.

Campbell differs from Popper in being more open to the idea that scientific development is guided by social selection. He wanted to go one step further and explain science's adaptation to reality as a relatively mechanical process in analogy with organisms' adaptation to the environment. However, it is unclear how far he thinks that social selection determines epistemic selection. He distinguished usefully between scientific justification and competition between scientific theories and hypotheses. The explanation of scientific justification is based on cognitive skills acquired by organic evolution, but when it comes to the rivalry between the alternative theories and assumptions, Campbell seeks an explanation in terms of social factors. Such social factors that act as selective forces include social recognition, persuasion, interests, peer pressure, institutions, economic resources, and social power; all these factors act together in the selection and creation of certain scientific traditions and views. Thus, according to Campbell, the diachronic understanding of science must be at least partially internal, because the individual observer chooses those assumptions, which have been best adapted to the environment by the logic of falsification, and the best adapted assumptions are those that have not yet been disproved.[5] At the same time the understanding of science must be partially an

[5] Campbell & Paller (1989).

external comprehension, because the social circumstances determine which competing ideas survive in the long term. It is therefore not surprising that Campbell seeks support for social selection theory in the modern sociology of knowledge program.

2.2 Setting the Legacy Straight

The strong empirical support for the Darwinian account of evolution indicates that the main idea behind evolutionary epistemology is correct; now the most rational thing to believe is that human cognitive capacities such as the principles of language learning and rational thinking are already embedded in us as "epigenetic rules"—an expression due to the sociobiologist Edward O. Wilson. Michael Ruse has commented that he feels less comfortable today in using this expression, because people invariably associated it with Wilson's theory of cultural development, so he prefers a more neutral term like "innate dispositions" or "capacities," and this is how we should read him.[6] Human cognitive abilities are "hardwired" dispositions encoded in our DNA. Obviously, some philosophers might choose to ignore this fact and still rely on a transcendental understanding of human cognition, but doing so flies in the face of everything our empirical research into human nature reveals. Experience tells us otherwise. Nonetheless, there are reasons to reflect on a weak version of evolutionary epistemology. Otherwise, the program may easily turn into a sort of scientific occultism.

Here it is necessary to distinguish between a weak and a strong evolutionary program. The *weak program* argues that cognitive mechanisms such as the individual's capacity of thinking, reasoning and language acquisition can and must be understood in terms of human evolution formed by selection and adaptation. However, the *strong program* goes further in that it also wants to understand collective phenomena such as society, science, and culture and their developments as a result of the same kind of selective processes.

Accordingly, Michael Bradie distinguishes between "The Evolution of Epistemological Mechanisms" (EEM) and "The Evolutionary Epistemology

[6] Cf. Ruse's Preface to the second edition of *Taking Darwin Seriously* ([1986]1998), p. xii.

of Theories" (EET).[7] The first kind of evolution applies to individual observers, where the cognitive selection begins with an individual, and then by virtue of a better fit of this particular individual's properties to the surroundings the cognitive mechanism is spread to its descendants and eventually to the entire population. In the second kind of evolution ideas, thoughts, and theories are assumed to evolve in relation to social circumstances. The selection of ideas, and thus a better adaptation of scientific theories to the social environment, is tied to a social fitness of these ideas in a community of ideas.[8] The strong program differs from the weak program in that it implies that selection is the principal explanatory principle behind any social or cultural development, whereas the weak form merely holds that selection explains the evolution of the cognitive abilities of individual organisms. The weak program allows the major development of social, cultural, and scientific ideas to be due to deliberation and decisions and not selection.[9] In other words, the weak program accounts for our *cognitive dispositions* and their function; however, the strong program also attempts to account, at least partially, for the actual *cognitive content* of these dispositions in terms of biological and/or social selection.

Whether we consider the weak or the strong version, it is important to keep things separated that are not always distinguished. Variation, selection, and adaptation can explain the emergence of skills, dispositions, and cognitive principles as genetically inherited characteristics of the individual. It does not claim that variation, selection, and adaptation can explain the knowledge we actually have or are able to get. The knowledge we actually

[7] Bradie (1989), p. 394. The distinction is important because early criticisms of evolutionary epistemology such as Thagard (1980) were not aware of it and were therefore successful in their rejection of the blind selection of theories but had little to say about a possible acceptance of the selection of human cognitive mechanisms.

[8] I think there are two versions here as well: A weaker one, which claims that in cases of more or less empirically equivalent theories, social factors determine which is chosen, and stronger version, which holds that all theories create their own empirical base, by virtue of the theory ladenness of observation and incommensurability of standards, so empirical evidence counts for naught (or nearly so), consequently the only basis on which theory choice is really made is social (or economic, or gender-based, or religious, or…).

[9] Among the supporters of the weak evolutionary epistemology one finds Michael Bradie himself, Quine, Wilfrid Sellars, Michael Ruse, Abner Shimony, Nicholas Rescher, Daniel Dennett, Eliot Sober, and many more, while those who subscribe to some form of strong evolutionary epistemology count names like Karl Popper, Donald Campbell, Stephen Toulmin, William C. Wimsatt, David Hull, and a number of others.

possess, and all we can ever attain, is *acquired* and therefore does not rely on genetic selection. A person's knowledge varies greatly from individual to individual, from culture to culture, and seems at least as much shaped by historical and social factors as by individual genetically inherited dispositions. Of course, the supporters of strong evolutionary epistemology may rightly point out that the difference between inherited and acquired characteristics is not so great. Innate cognitive capabilities are due to information that is genetically transmitted to the next generation, whereas acquired characteristics are due to information or skills that are linguistically and physically distributed to succeeding generations. Our use of language is an ability that has arisen through biological selection, because language (much more clearly better than other behavior) has shown its functional usefulness in conveying information between organisms in a different way than the purely genetic transmission of information. Consequently, the strong evolutionary epistemologist may argue that corresponding to this other form of information transmission there exists a selection mechanism that makes more likely that published, but individually supported, ideas and assumptions are adapted to the social and cultural structures that are grounded in language. However, the crucial point is whether the strong evolutionary epistemologist—in her eagerness to utilize 'selection' as an explanatory principle for cultural and social issues—has not easily come to miss a significant distinction. The distinction is between the role these innate capacities play in determining which ideas and assumptions we *possibly* can have and the role these capacities have in explaining why we as a collective *actually* have those ideas and assumptions that we do have and how we periodically change these ideas and assumptions.

So where do the boundaries between nature and culture become manifest, and does it make sense to appeal to the mechanism of selection when we talk about higher-level thinking and creative processes, of which the sciences are examples? How far can social norms, ethical standards, political ideologies, and scientific theories be said to have been at any given moment adapted to prevailing social and culture conditions through selection? To answer these questions we must first have a clear idea of what is included in the notion of "selection."

The word "selection" means that something has been sorted out. But this is a special form of assortment, where what has been selected does not have an influence on whether it is selected or deselected. In the case of natural selection it is the adaptation to the physical environment that determines whether an organism's descendants are selected or discarded. Although we can speak of natural selection as "favoring" those individuals who perform well, this is always a metaphor, for "to favor" implies conscious intent. In fact natural selection is always *passive* in virtue of nature's *passive* "sorting-out." If we look at epistemic selection, we encounter the same *passive* "sorting-out." This happens whenever empirical facts, according to Popper, passively "filter out" individual hypotheses that are not supported by observations or experiments or whenever social factors passively "filter out" socially accepted theories and models that do not suit the present social circumstances.

This evolutionary model is not compatible with our prima facie belief that scientists actively and deliberately choose their hypotheses. Unless such a notion is a fiction of our imagination, the strong program of evolutionary epistemology lacks an explanatory model in which scientists make an *active* choice of hypotheses in relation to the goals, which these hypotheses are constructed to help us to achieve. The proponents are being misled by the tendency to think of "Nature" as acting like a conscious organism. Natural selection is, of course, blind; it does not, so to speak, "know" what it is doing. But the scientist's selection of hypotheses is a conscious process and embodies a deliberately intentional rational choice of hypotheses and an acceptance of certain empirical facts.

Natural selection among a varying population explains genetically determined evolution. It is characterized by the fact that external information, i.e., not encoded in the genome, cannot improve or change the inherited properties of the organism. Darwin was correct as he, unlike Lamarck, denied that acquired traits could explain biological development. But at the same time it is reasonable to assume that at some point during biological evolution just by pure chance an inherited capacity arose, which enabled organisms with this capacity to receive external information which they could actively "choose" to accept and use constructively to build and improve the usefulness of already received information. In the end, this capacity of conscious selection made it possible for external, non-genetically transmitted information to be stored as learned skills, behaviors, knowledge, and values. Hence once the organism becomes capable of acquiring information by learning, it becomes possible for it to be enlightened by this information and to act upon it. In this way such information eventually became communicable from one generation to the next.

We must assume that biological development supplied us with a capacity to actively "choose" a particular conjecture, in light of the information we have received empirically about the relevant facts. Already long before hominids appeared on the African savannah, organisms learned to know their environment, or at least those parts of it that affected their survival and reproduction. It is inconceivable that this development can be understood without seeing the learning process as rooted in an inductive mechanism that must not itself be learned but must be transmitted genetically. A capacity of simple induction helps organisms to gain knowledge of their surroundings and helps them to conceptualize important things in their habitat. Of course Popper denied this. Thus at some time through a passive selection by the surrounding environment organisms must have gained a capacity of active selection. At some point in biological evolution, organisms that could store and accumulate their sensory experiences, gained a greater chance of survival, and thus the genetic information enabling the development and enlargement of memory was then passed down to their descendants. The capacity to build up experience, that is, to learn, is based on at least four or five key cognitive mechanisms: memory, concept formation, imagination, conjecture, and inductive inference. Indeed this capacity was strongly reinforced in human beings when the capacity for linguistic communication also evolved. Once these innate dispositions were in place, humans could actively begin to construct ideas and beliefs and shape hypotheses about the world that they cannot see. Put somewhat paradoxically: The selective explanation of human cognitive development also explains why selection is insufficient to explain the development of science.

First, nature passively opts out, and then the organism actively opts in. An active choice presupposes goals or intentions. How can we possibly explain the emergence of intention, intentional learning, etc., without having to refer to an already innate intention or design? The American philosopher Ron Amundson points out that a selective explanation must meet three key conditions: (1) There must be a multitude of variations, (2) variations must be of such a nature that they are random and not directed toward any goal, and (3) the sorting mechanism cannot be goal-oriented or purposive.[10] If a phenomenon does not meet these three conditions, we quite simply have no selective explanation for this phenomenon. When talking about phylogenetic development, it is precisely in relation to this natural evolution that Darwin's theory provides a convincing explanation in terms of natural selection. However, when we consider ontogenetic development the same sort of explanation seems unsatisfactory (at least when we are talking about human beings and other animals with a high-level awareness).

Nonetheless, certain forms of ethological behaviorism have sought to develop a selective model of animal and human learning founded in a stimulus-response mechanism. This approach assumes that an individual organism receives many different stimuli (information) from the environment, to which it responds randomly. Over time its reactions are gradually adjusted to the given stimuli through a reward (punishment) system to build up a positive or negative correlation between stimulus and response (conditioning). In other words such behaviorism subscribes to a kind of passive learning. The reward system consists of a mechanism by which the organism can learn to repeat certain behaviors since it gains some benefit from the environment.[11]

However learning is much more than conditioning. Three models of learning are needed to deal with different level of cognition. The genetic model of learning can handle examples like this: a termite queen can distinguish how much salivation workers and soldiers have left in the food with which it is fed and react to an imbalance by producing more eggs containing soldiers if necessary. At one point in the past the termite queen received information about an imbalance and she responded by starting to produce more soldiers than workers; eventually it was rewarded by receiving information that the balance between workers and soldiers was re-established and it began to produce fewer soldiers. A second model based on conditioning may explain an example like the following: many juvenile songbirds first learn to sing like adults when they have listen to

[10] Amundson (1989), p. 417.

[11] Of course the organism is still utilizing some sort of inductive mechanism: If an organism was rewarded when it responded with *R* last time, it will be rewarded when it responds with *R* next time. Conditioning assumes inductive capacities. So the behaviorist poly does not get you out of the need to assume that the organism has genetically acquired inductive capabilities.

the mature song from other males. Here it seems correct to assume that a sort of operant conditioning is at play. The behavior of the juvenile male bird is reinforced in virtue of having positive effects on the environment. This enables it to compete with other males.

A third model of learning may be called the intentional model, i.e., a purpose-oriented model of leaning. The problem with behaviorism and the model of conditioning is that it completely ignores the fact that higher animals are endowed with a reflective consciousness through the process of biological evolution. This means that under the influence of natural selection these organisms have developed an advanced consciousness capable of presenting the outside world and of imagining the expected consequences of their behavioral choices, given this presentation. The function of consciousness is to receive and process information about the environment where the organism's responses to this information are not already generically predetermined but depend on imagination and already stored information. This is a much more efficient way to process information that may be challenging for an individual organism, since the organism can respond more effectively by an individual conscious action rather than by a phylogenetic adaptive one. A reward system must have been in place long before such a capacity for reflective consciousness allowed for the possibility of individual conscious choice, but once the reflective consciousness evolved, the phylogenetically based reward system was supplemented with an ontogenetically based one.

Of course, the behaviorist may argue that these innate abilities function as internal constraints on the possible responses to external stimuli. But such restrictions reduce the possibility of selective explanations, cf. Amundson's condition (1). In addition, it seems fair to assume that consciousness, reflection, and language have had a survival value for individuals with these cognitive capacities, because they have contributed to an active individual learning.

Epiphenomenalism, which considers the mind as just an unproductive and non-functional property of the brain, has difficulty explaining the survival value of consciousness, just as it is impossible for epiphenomenalism to explain why consciousness has become increasingly developed as brains of our ancestors have grown larger and larger. For example, it is known from Hawaii—whose geological time is quite short—that cave-living insects and spiders, which have evolved from forms with colors and big eyes, have now turned white and have much reduced eyes. Eventually, their eyes will completely disappear. Animals that live in total darkness need neither colors for protection nor eyes for navigation. It is therefore biologically uneconomical to maintain these properties, especially because as they adapt to a life in the dark the affected organisms need energy to develop other sensory organs that can replace their eyes. An epiphenomenal consciousness without a function would also be biologically uneconomical. The mind helps a higher organism to act consistently with an active orientation relative to its surroundings in three ways: first by collecting and processing information concerning the present, second by deliberating the most likely future scenarios based on information about the past, and third by acting accordingly.

Variation and natural selection provide a convincing explanation of phylogenetic adaptation, but not one of ontogenetic adaptation. We may agree that the mind, as a capacity for presenting the environment, is created by variation and natural selection, but since natural selection is extremely slow relative to the lifespan of individual organisms and requires the presence of a large amount of variation in a given population to be successful, any mechanism that accelerates an individual's capacity for adapting to its environment would immediately have an obvious survival value. The *intentional* mind is precisely such a mechanism. A single organism that may actively and appropriately opt in and opt out in response to different information can much more quickly adapt to the actual environment than a species can adapt by passive natural selection. The alternative to a purpose-oriented adaptation by leaning would require a random "trial and error" learning mechanism, i.e., a mechanism by which the external circumstances sort out among many randomly generated guesses. The burnt child shuns the fire, they say. One time is enough. The probability that the individual quickly might get hurt, would be too great if (a) it could not use its past experience purposely as a deliberate means to solve new problems and if (b) it could not imagine what solutions were relevant and most effective. Purposeful solutions to the problems of everyday life are not chosen by chance; they are deliberately elected because they are likely to be more successful than arbitrarily chosen. We set up not just solutions indiscriminately, but choose possible solutions that we believe are relevant and have the greatest chance of success. This stands in contrast to Amundson's condition (2).

The basic difference between a purposeful understanding and a "trial and error" response can be illustrated by the following example. Imagine you receive an empty wine bottle where the cork is pushed into the bottle. At the same time you get a hammer, a metal hanger, a corkscrew, a spoon, a shoe, and a serviette. Now you are asked to get the cork out of the bottle without breaking it. Based on your previous experience with a hammer you immediately exclude it. The same reasoning probably applies to the spoon and shoe. You do not even try to use these tools. You note just by looking at the corkscrew it might be too short, or maybe you realize that a corkscrew cannot be used on a loose cork. Many will probably try out the hanger, making it a metal hook and try to fish out the plug. This seems a rational possibility until either a little experience or reasoning shows it will not work. If you are clever enough, then you do not even try to use the metal hanger. Reasoning by elimination eventually you realize the serviette is the tool you need to get the cork out. If one corner of the serviette is inserted into the bottle, and the bottle is turned upside down, the plug will lie between the serviette and the bottleneck and the cork can easily be pulled out. You solve a new problem not using random verification, but using experience and reflection.

Bearing in mind Amundson's condition (3), one may object that we normally do not reject the assumption just because it might seem to run counter to our immediate observation. It is not so that sensory experience serves as an automatic sorting mechanism among our assumptions. Often we must engage in some critical reflection on whether the "falsifying" observation is not what it purports to be, or whether the "falsified" assumption should be rejected, or which one of a conjunction of possibly falsified assumptions we choose to reject. Such a deliberative process does not meet Amundson's condition (3). In fact none of his three key requirements of selective explanation are met when we want to explain the outcome of reflective consciousness and advanced behavior.

Consequently, when we want to describe the ontogenetic development of a member of the human species, or other higher species, it is impossible to understand this process as purely a result of the Darwinian selective process according to which external circumstances constantly sort out among a diversity of individual randomly generated assumptions. Evolution has equipped us with the cognitive mechanisms that are capable of finding optimal solutions in the shortest possible time. It involves intentions, tactics, and strategies. Higher animals have been equipped with cognitive mechanisms that make our thinking, learning, and behavior goal-oriented and purposeful. In addition to them, the human species is equipped with a capacity of language that enables it to explain a world beyond our sensory capacities by the use of abstract concepts. These abilities and skills form the basis of the development of the sciences. Also in the sciences, we encounter reflection of the same nature as in everyday life. Despite the fact that Karl Popper held that science never uses inductive methods and believed instead that scientific progress consists in imposing bold guesses and then trying to shoot them down, there are many reasons for believing he was wrong. Falsification of scientific hypotheses happens all the time, but it is not an end in itself, or necessarily the end of the hypothesis. Thomas Kuhn gave an accurate criticism of Popper's falsificationism. There are, he said, only two alternatives: Scientific theories are never confronted with a counterexample, or they are confronted with counterexamples all the time.[12] For example, astronomers knew for many years that Mercury's observed perihelion movement differed from the one Newtonian mechanics predicted, but during the Newtonian era no one considered that falsifying evidence (for that is what it was) as refuting Newton's mechanics. Only after Newtonian mechanics gave way to relativity, which made the correct predictions, was this evidence seen as falsifying Newtonian mechanics.

Science is progressive; biological selection is not. In contrast to biological change, which is random [and reflects utterly unpredictable environmental change for innumerable non-biological reasons (geological, meteorological, astronomical, pathological, etc.)], scientific change, like all cultural change (of which it is a species) in general, builds deliberately upon earlier ideas. So it seems. Sometimes the lack of progress in biology is questioned due to the fact that humans are more complex than lizards, which again are

[12] Kuhn ([1962]1969), p. 80.

more complex than worms.[13] So even if there is no teleological progress in nature, there seems to be a direction toward greater and greater complexity. How can random selection produce such an improvement? Understanding this controversy hinges on what we mean by complexity, but in order to make the discussion short, let us assume that complexity has something to do with structure: How many elements a system contains and how many causal relations that exist between these elements indicate the degree of its complexity. Given enough genetic variations in a population, some will produce an offspring with a trait that has a bit more complex structure than its parents, some of them not. Whether one of the more complex forms is selected or not selected depends on the environment. There is no guarantee that a more complex form is more adaptive than a simpler form, but sometimes the complex variation pays off, in which case the more complex form eventually will dominate in the population. Accidentally, evolution took a direction, which has now produced human beings.

However, in self-conscious beings the introduction of new ideas is goal-orientated. New ideas relate to what has been the tradition and modify or change it in a purposeful way. Where biology is Darwinian, culture is 'Larmarckian' or 'Spencerian.'[14] As true as it is that science and culture transcend biology, they still have their roots in biology. The Darwinian approach to epistemology must be able to explain the 'Larmarckian' nature of cultural change. Natural selection from random variations explains the development of the *capacities* of human thinking and the *forms* of thoughts, but the *content* of human thinking is somehow determined by reflection and conscious intentions in the meeting between the physical and social worlds. Of course, it is imperative for the Darwinian to explicate how random changes among the genotypes can cause non-random change among the phenotypes. Two strategies seem available.

Michael Ruse points to the distinction between what he calls the environmentalist option and what might be called the epigenesist option.[15] The first takes the human mind as *a tabula rasa*. Whatever there is to

[13] For a critical discussion of progress in biology from a Darwinian point of view, see Ruse (2012), pp.99-127.

[14] See, for instance, Ruse ([1986]1998), pp. 124–125.

[15] Ibid., p. 140–141.

learn, humans can learn it. After ideas have surfaced in the human mind irrespective of their adaptive value, the environment starts to annihilate the maladapted ones. This is the trial and error-option we have already discarded. This method is inefficient, inter alia because, as Ruse says, humans have "to learn all sorts of things in each generation which could as readily be passed along encoded within the genotype."[16] Moreover, the environmentalist option hardly takes us beyond Darwinism. Humans do not get to their conclusion by reflection and deliberation. In every new generation it is the environment alone that determines what we believe and what not, and for those whose beliefs are unfitted it may have fatal consequences. In contrast, the epigenesist choice denies that the human mind is a *tabula rasa*. Instead this view holds that our way of thinking and doing things is deeply embedded in our genes, since it is biologically advantageous for us to think and do these things in particular ways. If our beliefs are already predetermined to a high degree by nature, individuals can think and act within the bounds nature makes possible without devastating results for their survival or reproduction. Here Ruse points to Wilson's epigenetic rules (innate dispositions). These are inherited as cognitive matrices in the human genotype and generate innate reactions to external information under the right circumstances. However, some of these matrices seem not only to shape the form of our thought and action but also fix their content. Nevertheless, "epigenetic rules" cannot yield the complete answer because human thoughts and actions would then be absolutely determined by our genotype. The fitness of an organism is much more benefitted if evolution likewise provides it with epigenes that code for certain cognitive capacities allowing the phenotype to choose the content of its beliefs in the form of imagination and reflection. As a result we see the flourishing of human cultures in all their shapes and colors.

Therefore I conclude that evolutionary epistemology in its strong form is untenable. Selection cannot explain the development of science. However, the weak form seems to give us a successful explanation of why we have the innate and goal-oriented skills we do have. An explanation for the development of the sciences should be sought not in a process of social or cultural selection, but in the interaction between our innate

[16] Ibid., p. 142.

ability of imagination and the surrounding reality. In addition to the way we understand and represent things, natural selection has also given us a generically inherited capacity to judge whether what we think and believe also corresponds to that which we think and believe. Biological evolution by natural selection has provided us with a facility to consciously troubleshoot false assumptions and maintain true beliefs.

2.3 A Fallacy of Naturalization

While the foregoing considerations seem to rule out the strong program, there are more far-reaching arguments against evolutionary epistemology, which, if sound, cast doubt on not only the strong, but also the weak program. Traditional epistemology is a normative discipline. Evolutionary epistemology is descriptive. Traditional epistemology therefore seems to explain something about human epistemic behavior that cannot be explained by a purely descriptive, naturalistic theory. A normative theory can tell us that a person acts in an epistemically responsible way, if and only if he or she feels obliged to uphold or reject a belief based on certain criteria of good science. Therefore many will be skeptical of whether or not a descriptive discipline such as evolutionary epistemology can possibly address the normative goal of traditional epistemology.

Popper, who supported evolutionary epistemology, believed that his EET model of falsification of scientific theories provided a normative methodology.[17] His problem was that a descriptive theory such as Darwinian evolution cannot also be normative—unless, of course, one commits the naturalistic fallacy and infers from what is to what ought to be. It is also not a good idea, one may argue, to tie epistemology closely so a scientific theory such as natural selection. Since all scientific theories have the potential of being abandoned, therefore the descriptive element needs to be qualified. However, the falsification of the descriptive element does not imply that the normative element will not fall and therefore needs no modification. For instance, Bradie considers the normative element of epistemology as the one that lasts, since this is not based on facts in the

[17] See Popper (1972).

same way as the descriptive element. According to him, Popper is not in a position to argue that falsificationism sets out a framework for how the researcher ought to act and at the same time describes how nature itself actually disposes of wrong ideas through falsification.[18]

As a whole Bradie supports weak evolutionary epistemology. But he also recognizes that the position faces grave difficulties in answering a number of fundamental epistemic questions. In particular he points to three urgent snags:

The first challenge relates to the naturalistic fallacy.[19] Discussing this fallacy Bradie confronts Hilary Kornblith's view. Kornblith argues to his own consent that one subscribes to a naturalized epistemology only if one answers in the affirmative to the last of these three questions:

Q1: How *ought* we to arrive at our beliefs?
Q2: How *do* we arrive at our beliefs?
Q3: Are the processes whereby we *do* arrive at our beliefs *the same as* the processes by which we *ought* to arrive at our beliefs?

Traditional epistemology emphasizes that Q1 and Q2 can be answered independently. Therefore Q3 is denied. Any naturalized epistemology, of which evolutionary epistemology is a species, holds the opposite view, namely that they cannot be understood or answered independently. Critics of naturalized epistemology would reject this claim on the grounds that such a view commits the naturalistic fallacy, because the answer to Q1 is derived from Q2. Apparently, Kornblith has two options: he can either argue that the naturalistic fallacy is no fallacy, or he may say that Q1 cannot be answered independently of Q2. Kornblith himself prefers the last alternative. And he adds that when Q2 first has been answered, then there is nothing more to say about Q1.[20] That's what he calls *the replacement thesis*. This view provokes Bradie to say that thus Kornblith commits the naturalistic fallacy with a vengeance.

18 Bradie (1989), pp. 408–409.

19 Ibid., pp. 395–396.

20 I presume that this is what Kornblith actually takes to be Quine's view as it is presented in "Epistemology Naturalized," i.e. that epistemology must abandon its traditional normative aspirations; saying what we "should do" is nothing more than saying what we actually do.

If one accepts Bradie's interpretation, it becomes problematic for the advocate of evolutionary epistemology to answer key epistemological issues, as they are mainly normative, for example, the issue of entitlement to a belief. Can a purely descriptive approach answer normative questions without committing the naturalistic fallacy? Bradie thinks not, arguing that the relative independence between normative and descriptive matter should be maintained. He admits that he sees no immediate solution to the problem.

Unlike Bradie, I am much more optimistic about the naturalization of epistemology. Norms are undoubtedly man-made. Nevertheless, they are based upon the abilities that we actually have. It makes no sense to say that one ought to do something unless it is assumed that one can do it. Of course we would not have a norm saying that it is 'normally' wrong to tell a lie, if it were not because we actually have the ability to lie and the ability to speak the truth. The capacity to assert what is false, i.e., to lie, is innate; it exists already in *corvids* and higher mammals and, of course, is given to those who have utilized deception for some survival advantage. But to keep someone deliberately in the dark, one must be able to imagine what actually the case is, and what would happen if one did not act deceptively, and what would happen if one did. One must be able to represent the environment and find out what the consequences are if one does such and such. Such capabilities are directed toward the achievement of a goal, although their origin is based on natural selection. Such capabilities, which work by finding out what is most appropriate for a particular purpose, also provide the ground for the development of more reflective cognitive systems, where one has the opportunity to consciously formulate epistemic and moral standards with respect to general man-made principles. What matters here is that when the capacity to imagine an object first appeared, it was no longer a selective process that determined which purposes one actually imagined an action should have. Such purposes may aim at something actual, something constructed, or something highly imaginary. One purpose may be to think and reason properly, another might be to do things right.

Since traditional epistemology defines knowledge as justified true beliefs, it also involves defining standards for what it takes for a belief to obtain a status of knowledge. Beliefs must be justified. Therefore the internalist demands that one needs to know that one's beliefs meet the standards of justification. The naturalist does not have to deny the requirement that scientific beliefs should at least be empirically justifiable and that on the basis of induction, abduction, and inference to the best explanation we have the possibility of showing whether or not our beliefs meet the requirements we set for justification. The purpose of justification is precisely to show that a particular belief is supported by norms that are considered to lead to genuine knowledge. If these considerations are on target, the naturalistic fallacy is a problem only for the strong form of evolutionary epistemology. Weak evolutionary epistemology can give a good explanation of how the normative sides of cognition gradually emerged.

The second challenge, which Bradie discusses, involves a vicious cycle, a problem initially raised by Roderick Chisholm.[21] The vicious circle is based on two fundamental epistemological questions: (a) what do we know? (the problem of extension) and (b) what are our criteria of knowledge? (the problem of criteria). Is it legitimate to use scientific knowledge in our epistemological interpretation if we afterwards use the same interpretations to evaluate scientific knowledge? In the light of this challenge Chisholm distinguishes between two epistemological approaches. One is *methodism*, according to which we should formulate an answer to (b), i.e., to adopt a method first, and then use it to respond to (a). The other is *particularism*, according to which we should formulate a response to (a), giving model cases of genuine knowledge and then use it to respond to (b).[22] John Locke and other empiricists are methodists, while naturalized epistemologists like Popper and Campbell are particularists. Bradie points out that the circularity not only perpetrates the naturalized epistemology variant but any theory of knowledge. At the same time he doubts that Darwinian thinking can defend the methodist approach over the particularist.

But is the allegedly vicious circle really vicious? What we know and what the criteria are for knowledge can easily be established through a gradual specification of different levels of knowledge. At first, we can establish a purely functional sense of knowledge. Like humans, many animals possess knowledge that originates in their capacity to store and use sensory information from the environment to behave successfully. A guillemot finding

[21] Bradie (1989), pp. 401–403.
[22] Chisholm (1982), pp. 61–72.

its nest among tens of thousands of other nesting sites on the bird cliff has such knowledge. Indeed this is an embodied and practical form of knowledge, but it is *acquired* knowledge nonetheless. First in those cases where we talk about propositional and reflective knowledge, gleaned from the information we receive directly from our senses, do we need to establish certain criteria for proper knowledge acquisition. It is necessary to have criteria of justification only in these situations, since under these circumstances the process of mental reflection occurs in order to bridge the conceptual and inferential gap between what we know directly and what we might possibly know indirectly by inference from the directly known. In order to bridge this gap, one requirement for satisfactory justification is that it be rooted in an adequate empirical basis. One can then appeal to normative criteria of adequacy to determine which of our many reflectively acquired conjectures meet this requirement. The same holds for other requirements.

Finally, as the third challenge, Bradie focuses on what he calls the Darwinian argument for naturalizing epistemology.[23] His starting point is Kornblith's claim that truth has a survival value and that organisms not having cognitive contact with the world have a pathetic but praiseworthy tendency to die out before they reproduce.[24] In Bradie's reconstruction the argument goes like this: (1) The belief in truths has survival value. (2) Natural selection ensures that our innate intellectual endowment gives us the capacity to believe in truths. (3) Knowledge is a necessary byproduct of natural selection. Kornblith finds support for this line of arguments in Quine and adds the following in support of the replacement thesis: (4) If nature has designed us so that we tend to prefer true beliefs, then the processes by which we arrive at these true beliefs are the processes we ought to use. Hence those processes by which we arrive at our beliefs are those we should use. The conclusion can be formulated in a strong and a weak version: The strong variant holds that there is complete overlap between how we actually do reach our beliefs and how we ought to arrive at them, and the weak variant claims that there is only approximately a

[23] Ibid., pp. 403–406.

[24] Kornblith (1985), p. 4 f. Of course we talk this way semi-seriously, but strictly speaking such organisms never "die out" because natural selection would never have allowed a *species* of such creatures to have evolved in the first place. *Individuals* cursed with "lying" faculties just die, so no such species could exist.

match between how we reach and how we ought to arrive at our beliefs. It is important to keep in mind the difference between being genetically predisposed to find truth and holding assumptions that are in fact true. Kornblith's argument says only that we are predisposed to find truths. But then the second part of the argument states that the methods and cognitive mechanisms we actually use when we pursue our natural dispositions are also the ones we ought to use. Furthermore, we might add the assumption that the methods and cognitive mechanisms that we actually use to find truths must also be the ones that we are predisposed by selection to use. But from here we cannot reach the conclusion that they are the same methods that we ought to use, because if we are predisposed to use them, we can hardly fail to use them. Thus, it does not make sense to add the normative claim that we ought to use them, since we cannot do otherwise. So the naturalistic fallacy still blocks the way to a normative conclusion. Are those methods we are predisposed to use also those we ought to use? One possible riposte is to ask another question: How can one be free to choose a method we ought to follow if natural selection has determined it otherwise?

Bradie's criticism is different: he argues (2) does not follow from (1). The mere fact that some characteristic of an organism has survival value does not guarantee that it will be selected in the selection process. Therefore, (3) does not follow from (2). Bradie's point is true, of course, but Kornblith's assumption (2) does not claim that there are no properties that would have had survival value for human beings, even if these properties never have been or will be realized by a biological selection. There are probably many such properties. Kornblith's claim is rather that the selection process has actually selected those organisms that were predisposed to receive information that does not misinform them about their surroundings, that is, were actually genetically predisposed to form true beliefs rather than false ones.

According to Bradie, the conclusion (3) is in itself problematic. The argument indicates, he claims, that evolution ensures that organisms and their environment "fit together," but it says nothing about the process by which humans acquire knowledge, much less if the beliefs they do acquire are genuine knowledge (as traditionally defined). Assuming that our cognitive abilities have evolved by natural selection does not lead to the conclusion that the mechanisms that are alleged to generate knowledge also are predisposed to generate truths rather than falsehoods. Obviously, Bradie is right that biological evolution is not a good model for the development of scientific knowledge. It is possible to recognize the importance of Darwinian selection in the formation of the basic cognitive mechanisms without underwriting the further belief that science follows the model of selection. But I believe he is wrong in claiming that the very basic cognitive mechanisms that provide us with empirical evidence are not better adapted to the environment than that they are not more likely to endow conscious organisms with true beliefs rather than false beliefs. If these knowledge-generating mechanisms worked quite randomly and induced ignorance as often as knowledge, they would soon have a negative effect on the survival of the genotype. The organism has no other means for determining if information from its senses is reliable, and thus its behavior could not be designed appropriately for the environment. For the same reason it would be inexplicable why natural selection would not have selected against such ignorance forming mechanisms, but instead seems to have been reinforced in virtue of the continuing phylogenetic development of increasingly complex cognitive structures.

Thus, I conclude that the weak version of evolutionary epistemology can avoid committing the naturalistic fallacy. The argument goes as follows: the innate cognitive mechanisms that we actually possess, and which selection has predisposed us to use, have proven effective for their selected function, namely more frequently to provide us true than false beliefs. Of course, this does not allow us to conclude that it is also the mechanisms we ought to use. Such an insight can only be reached after *Homo sapiens* had developed a capacity for reflective thought, which can assess its own cognitive mechanisms with respect to self-adopted objectives. With the capacity for reflective thought, which we have inherited from our ancestors, follows the ability to use our innate cognitive mechanisms purposefully and to evaluate them with standards of the mind's own adoption. Of course, we cannot come to the realization that our innate mechanisms ought not to be used. There is no alternative. It makes no sense in this context to talk about an "ought" when there are no other options. In contrast, with respect to reflectively acquired beliefs we may set up cognitive standards that we rationally believe must be met before such theoretical assumptions rightly can be said to express knowledge. In this situation we can talk about "ought" because, as reflectively conscious beings, we can easily intend our beliefs to adhere to various standards. But an ought follows a purpose, and its relevance is based on the success by which it helps us to realize this purpose.

Intentional behavior distinguishes higher animals from lower. First when the natural capacity for having intentions had evolved through millions of years of selection, organisms were able to interact with their environment by deliberate actions instead of just passively responding to it. Innate intentional behavior gave rise the construction of concept formation in many animals, and this was the beginning of the evolution toward capacity for language formation in humans. Especially with the development of language, the human species was able to create social and cultural communities. But here too, the capacity for language serves as a reliable means of communication, because mastering linguistic expression gave our ancestors an evolutionary advantage in comparison to those that relied on mere bodily gestures.

2.4 Intention and Innate Dispositions

One section before the last we touched upon "epigenetic rules," or better innate dispositions. It is now time for me to focus on how the methodology of science and scientific reasoning fits into this Darwinian picture of human knowledge. From what has been said so far it is obvious that the roots of scientific practice have to be found in the biological advantage science-making has for human reproductive success. Here I am not alone. Adaptation by selection is what justifies and constrains this practice. But I have also argued that science is a result of human intentions, powers of imagination, and reflective reasoning. The content and progress of science seems to be a product of human intellectual abilities to pick up external information and choose between different theoretical alternatives in order to explain that information. The conundrum is to reconcile these two assertions.

Epigenetic rules are characterized as "ultimately genetic in basis, in the sense that their particular nature depends on the DNA developmental blueprint… In cognitive development, the epigenetic rules are expressed in any one of the many processes of perception and cognition."[25] These rules have evolved as adaptations because of their survival value. Wilson distinguishes between "primary rules" and "secondary rules." Primary rules concern how raw information coming into the (human) organism is prepared for recognition; the secondary ones are those that go further to process this information in a way that is adaptively useful for learning, thinking, acting, and communicating. Although—like others—I shall continue to use his terminology, I shall change it a bit and add a further distinction. The "primary rules" correspond more or less to what I call "cognitive schemata", and the "secondary rules" are what I call "cognitive mechanism". "Epigenetic rules" as either schemata or mechanisms have their origin in the genes, but their status as cognitive and methodological *rules* results from abstraction due to human introspective reflection on the pre-existing cognitive practice itself. The objectives for such an abstraction are innate cognitive schemata or mechanisms, inbuilt in our way of thinking long before the human species had a language capable of expressing those rules. Some of these cognitive dispositions must at least go back to a time before the dinosaurs. But at some point in the evolution of humans or human-like beings the ability to become aware of one's cognitive practice was selectively established and indeed this helped their descendants to improve their cognitive practice. So I speak about "epigenetic rules" as the reflective counterpart to the existence of innate cognitive schemata and mechanisms.[26]

In the last three or four decades evolutionary science has generated overwhelming evidence that both the so-called "primary and secondary epigenetic rules" have a genetic origin. This evidence stems from many different sources such as transcultural experiments, pre-language experiments with infants, and comparative studies in relation to experiments with chimpanzees, dolphins, monkeys, rats, ravens, and other intelligent creatures. Basically the most successful hypothesis for explaining all these data is that many patterns of human reasoning are innate and not learned. In these experiments neither humans nor animals were trained to think in certain

[25] Lumsden & Wilson (1981).

[26] When did this reflective self-awareness begin? Was it after language or before? It would seem that the conscious formulation of abstract rules—as distinct from their concrete embodiments—could only have happened after both language and some sort of culture prevailed, so very recently in evolutionary terms. Then it would seem that cultural evolution rather than biological evolution would have come into play in establishing these schemas as rules.

ways; thus if they are able to solve certain tasks, regarding which they had no prior experience, the only reasonable conclusion is that somehow they had genetically evolved schemata of recognition and mechanisms of thinking that guided them to the goal. The mechanisms cover all kinds of reasoning: deduction, forms of induction, comparison, abstraction, etc. All these forms of recognition and reasoning are epigenetic rules brought into existence as cognitive schemata or mechanisms by natural selection. Nobody will deny, of course, that the human capacity of learning by self-reflection has enhanced the use of these dispositions far beyond what any other animal on earth is able to employ. This also means that most naturalists will agree that logic, mathematics, and the empirical sciences deal with issues that have no direct relations to our survival and reproduction. Nevertheless these sciences are constructed and receive their justification in accordance with these naturally evolved epistemic rules.

Central to reflection is the ability to target our action and thinking in the direction we want in order to fulfill wishes, needs, and goals. This capacity is what we call intention. It is itself grounded in a cognitive schema. By all accounts, having intentions is a feature human beings share with higher animals.[27] It is not a trait that just appeared when natural selection produced the first *Australopithecus* or the genus *Homo*. Mammals such as dogs, pigs, dolphins, and chimpanzees possess mental states, which may be characterized as having intentions. In short, an intention is characterized as a mental state that is directed at a target and acts as a cause for the behavior taken to achieve this goal. Such an intention does not have to be an object of awareness, or for that matter could even be made an object of awareness. It is not a requirement for being in an intentional state that the possessor is conscious or aware of its presence. Not many animals other than humans have the ability to be aware of their own states of mind by being able to represent to themselves their own mental states and therefore their own intentional states. So from an evolutionary standpoint reflective awareness presupposes

[27] *The Cambridge Declaration of Consciousness*, signed by many prominent scientists in the field of cognitive neuroscience and animal behavior, states in its conclusion: "The absence of a neocortex does not appear to preclude an organism from experiencing affective states. Convergent evidence indicates that non-human animals have the neuroanatomical, neurochemical, and neurophysiological substrates of conscious states along with the capacity to exhibit intentional behaviors."

pre-existing intentional states. The ability to reflect on their own mental states, including desires and intentions, has given *Homo sapiens* a great evolutionary advantage because it allows us to plan for the future and to change our natural surroundings.

Since we do not need to be conscious of our own intentions to have them, it is sufficient for being a mental state that it can cause achieving a particular goal toward which the intention is directed. The conscious intention is therefore a mental representation of a specific goal that one wants one's actions to realize and that in the right circumstances induces this action. This feature distinguishes intentions from other mental states. Usually it is argued that all mental states such as beliefs, feelings, and desires are "directed toward an object" in the sense that they represent their object, so the mind's directedness is also called its intentionality. But unlike the mind's other representative states, intentions function at the same time as causes of action, or at least have the potential to do so.

Explaining the diversity of plants and animals on a naturalistic ground, Darwin hardly regarded it as possible to find such an explanation outside of biology. He rather looked for a way to describe a uniform mechanism by which traits of plants and animals could evolve and which was common to all biological organisms. He focused on reproductive capacity, (genetic) variation, adaptation, and habitat. This gave him the opportunity to identify the causal mechanism behind evolution, which he called natural selection. But when it comes to the explanation of individual human behavior and attitudes, then the Darwinian explanation of the evolution of species is less applicable, if it is applicable at all. Indeed, one can often point to instincts as innate dispositions that are encoded in the genetic materials of various animals. This may be the birds' migration patterns, predators' hunting habits, or the various species' mating behavior. But on top of the instinctive behavior an intentionally controlled behavior gradually develops that differs from instinctive behavior by not characterizing the population in general but the particular individual. Intentions offer the individual the opportunity to put its own unique stamp on its behavior and, at best, open up many more behavioral possibilities in contrast to its instinctive, species-driven behavior.

The same is equally true of humans, even more so. Thus what the sciences are dealing with especially is the result of intentionally controlled behavior. From the point of view of the agent's own consciousness intentions appear, whenever we are aware of them, as internally given reasons for actions. However, from an external point of view they function as causes that produce the appropriate action by their ability to represent the effect in advance of its actual occurrence. If it were not for intentions, humans would not have developed language, tools, science, social institutions, and culture. We would not have created rules, norms, or standards. And none of us would have been able to improve our bad manners.

3

Sensation, Perception, and Observation

Our senses put us in touch with the surrounding world. In this empirical world of common-sense experience, things have extension, textures, colors, tastes, odors, and produce sounds. Evolution has formed it this way, and genetic variation and the environment have selected the specific functions of the senses through which we experience these qualities. Of course, we already know from many examples of animals that they have sensory channels vastly different from ours, so presumably what these animals "know" of their environment is quite different from what we know. But without the sensory channels mentioned we humans would be in no position to know physical objects or to interact with nature, nor would we have had any language. In the past the collective sensory experience of our ancestors led to the construction of our common-sense world. Even today it continues to constitute the appearance of the actually existing world, and our natural language is adapted to that fact. Sensation is our inborn manner of getting information about our environment. On the basis of beliefs acquired by sense perception, we can intervene in the course of nature, create new things, destroy them, transform them into something different, or chase them around.

J. Faye, *Experience and Beyond*,
DOI 10.1007/978-3-319-31077-0_3

In this chapter I distinguish among *sensation*, *perception,* and *observation*. The perceptual world is the visible world conceptualized, and the observational world is the perceptual world acknowledged. The growth of science uses technology to augment our senses and thereby enables us to take scientific knowledge beyond the world of direct sensory perception to the world of the unseen. The main message of this chapter is that what cannot be seen may yet be "observed" in the sense now used by the sciences. Thus, observation plays a very important role in the development of the scientific conception of the world and in the establishment of an experimental and a linguistic practice in science. But telling what kind of role it plays, and how it plays it, demands that some well-known controversies be addressed: What more precisely do we understand by observation; what is it to observe? What does it mean to say that two persons see the same thing? Similarly, there is the question of modality: what can we perceive when we observe? What kinds of entities are observable? And finally there is the question: How do we pick out those terms, predicates or sentences, which we characterize as observational? An answer to some of these queries will give us some clues to how closely our understanding of the world is bounded by our biology.

Perception is sometimes divided into two aspects: the phenomenological (or aesthetic) side of perception concerning the sensory form in which the content is presented to us and the propositional side of perception concerning the beliefs arising from the informational content of perception. I believe that historically, philosophers have emphasized the importance of the phenomenological element in their account of perception, at the expense of the propositional element. In doing so they have forgotten that perception is much more than merely experiencing certain sensations. More recently the focus has shifted to the propositional part: namely that perception is the acquisition of beliefs about the external world. However, cognitive science and evolutionary psychology seem to establish that perception operates on more than one level. So what is perception?

3.1 Perception as Belief Acquisition

Seeing allows us to gain non-inferential beliefs about the world surrounding us by means of our unaided eyes. It is common among philosophers and cognitive scientists to distinguish between sensation and perception. Sensation is the immediate sensory impression that informs us about our nearby surroundings, whereas perception is seen as interpreted sensory impression.[1] I think the latter characterization is unfortunate. Interpretation and cognitive processing are two different mental processes. Interpretation is a conscious process that consists of creating a hypothesis about meaning, and therefore involves forms of inferences, but what is called "interpretations of sensation" refers to cognitive processes that happen automatically and are not accessible to introspection.[2] The cognitive processing of sensory information consists in a spontaneous non-conscious transformation of this information into conceptually mediated beliefs. I think Sellars was on point when he argued that we have beliefs that we have acquired non-inferentially directly through the senses[3]: What he meant by this was that perceptual beliefs do not arise from the alleged interpretation of 'sense-data' or sense impressions. But he was so enmeshed in analytic philosophy that he thought that it was impossible to have any belief that did not presuppose other beliefs. However, in this respect, Sellars was mistaken; experimental evidence indicates that we can have the capacity of forming innate beliefs, which are not linguistically expressed and which do not require the existence of other beliefs.[4]

It is not only humans who can see. Higher animals, such as mammals and birds, as well as lower animals, like reptiles and insects, also have

---

[1] See, for instance, Gärdenfors (2006), p. 25.

[2] See Faye (2012) pp. 96–102 for a discussion of interpretation in the humanities and the postmodern claim that everything is interpretation. I believe that interpretation takes place in case we do not understand what we perceive. A scientist perceives the spectrum from a galaxy but she thinks it is abnormal in comparison with spectra from other galaxies. This forces her to make an interpretation of the abnormal one.

[3] Sellars ([1956]1997), sect. 32.

[4] For instance, it is evident that birds, fish, and apes have the ability to perceive colors. Thy need not be aware of their color beliefs to have them. Not even in principle. Awareness of the colors is enough to have color beliefs.

vision. All animals with sense organs have sensory impressions because they have developed a natural capacity to select and discriminate among things in their neighborhood in order to mate and find food. When they experience these sensations these animals are brought into a certain *sensory* state. But, although most animals do see in virtue of having visual receptors, nothing indicates that most of them are endowed with a belief of what it is that they are seeing. Seeing does not bring them into a *perceptual* state. But who actually knows whether or not the hawk perceives the lizard as a lizard; whether or not the mouse sees the cat as its enemy; whether or not the grizzly bear classifies the salmon as food; or if the he-wolf senses the she-wolf as a female? Of course, various animals behave differently toward different kinds of animals; but is it reasonable to assume that they do not recognize any particular animal as an object of a certain kind?

Although animals may seem mute it seems wrong in some sense to say that they do not hold beliefs. Indeed they do not have the power to express their beliefs linguistically, but, nevertheless, the bear seems to believe something is food whenever it recognizes it as edible, and the he-wolf seems to believe the she-wolf is a female whenever he mates with her. The bear can sense that different kinds of fruits and berries belong to the same category of edible things though these are perceptually dissimilar. Therefore the category "edible things" has a certain kind of perceptual reality for the bear. We could also include examples of domesticated animals whose behavior is to us most familiar: the dog surely believes one person is a friend and another a foe. They believe they will be taken for a walk every morning, then get fed, etc. They learn all this from humans, of course, but they are surely able to communicate the content of their beliefs in their behavior. In general one might argue that since many higher animals form communities with a well-defined social structure, and since they are able to display a great variety of behaviors, these phenomena can only take place if we recognize that these animals hold a variety of beliefs, which enable us to explain their behaviors. Classifying particular animals into groups of different species involves abstract thinking. Mammals and birds have the ability to identify their mates and offspring and even their own species; they can discriminate between what is edible and what is not and who are their predators and who are not. Thus, the argument concludes, this gives some warrant to ascribing beliefs to animals.

We are so used to thinking that concepts and language go hand in hand that it seems out of the question that animals use any kind of concepts. We learn the use of concepts by learning the use of a language. I suggest, however, that whenever an animal by its behavior show signs that it senses a token as a type, it masters a concept. So a belief arises in its mind whenever it actually recognizes a particular thing it experiences as a certain particular individual, such as its mating partner or offspring, or in those cases in which it recognizes a group of individuals to be particular instances of a certain type. I think there is good evidence for such a claim.[5] Hence, the ability of concept formation is an innate ability that found its way into biological evolution a long time before the appearance of *Homo sapiens*. Some concepts are inherent as innate dispositions; others are learned by imitation, imagination, and reflection. Indeed more advanced and sophisticated concepts used in human thought would be impossible without the use of *human* language.[6]

When it comes to animals' identification of other animals as belonging to their own species or their identification of other animals as 'prey' or 'enemies,' we find that even the lowest animals are able to pick out a mating partner and ferret out some suitable food and avoid being some other animal's food. Nobody would suggest, I imagine, that a spider, by making a visual identification of a mating partner, is acquiring a belief that it is confronted with a specimen of the opposite sex or that the honeybee has certain beliefs about where nectar can be found when dancing in front of the beehive. Why should it then be necessary for birds and mammals to have beliefs about what they are seeing in order to do what they are doing? The answer could be that beliefs come in degrees. Beliefs may vary from the very rudimentary to the highly complex through the

[5] The evidence is abundant. Each time an animal recognizes its own species it shows that it possesses a concept of that species. Since even lower animals can identify their own species automatically, it seems reasonable to assume that some concepts are inborn, whereas others are learned by experience.

[6] The issue about "language" depends in part on how we define "language." It is unquestionable that animals *communicate* with each other by a variety of symbolic systems that may use auditory techniques like whales and some birds, or use behavioural patterns, like building a nest or a bee's "dancing." They also "mark" various things in their environment as a sign of their presence to other animals. In all of these cases a semiotic system is used to communicate a message. Is that "having a language"?

various stages of the evolution. I hardly intend to claim that the spider or the honeybee is in any state close to understanding what it is doing.

Whether it is reasonable to attribute any beliefs to birds and mammals is a question of whether or not one interprets their behavioral accomplishment as demanding that their perceptions can be explained only by attributing them certain beliefs. I see no compelling argument against such an interpretation. I understand 'beliefs' to be mentally 'converted' information that *x* is *F*, either received or put into storage. Information is carried by a signal *S*, whose pattern an organism has the ability to process and respond to, so that the processing can result in a conscious or unconscious use of this information. If information is transmitted by a signal, then by virtue of that fact it is informative, it is not encrypted and so does not have to be 'decoded.' Rather in order for information to give rise to a belief, it must be converted by the brain in such a way that it becomes a mental property in the form of a specific belief about the world, giving rise to the signal *S*. However, the organism may process *S* incorrectly. It also means that if an organism believes that *x* is *H*, this belief may be false either because the creature has processed the information incorrectly or because the information content was deficient with regards to whether *x* is in fact *F* or *x* is *H*. Someone could argue, of course, that the question of whether dumb animals can possibly have beliefs hangs on the question of whether it is possible to have beliefs that are not conscious in the mind of the believer. To exclude such a possibility one can claim by definition that to have a belief is to have the capacity of knowing one's belief and making judgments about its content. But such a definition seems to be an act of fiat since it would also exclude any explanation of human behavior in terms of beliefs of which the agent is not aware. Although the question of unconscious motivation carries with it its own philosophical problems, there seems little doubt that even human motivation does not lie wholly within the realm of our conscious beliefs, much less what motivates other conscious or unconscious organisms in our environment.

A more evolutionary approach is needed. In my opinion it makes sense to attribute some beliefs to higher animals to explain their behavior. Even though these animals cannot deliberately justify their beliefs, due to evolutionary adaptation such beliefs have survival value for those animals. This suggestion fits nicely into the Darwinian outlook that consciousness is not a property solely of *Homo sapiens*. In general higher animals are involved in some kind of epistemic perception. The reason why it is true that animals other than human beings may not be aware of their own beliefs may have to do with their lack of a reflective consciousness.

Fred Dretske has distinguished between *non-epistemic* seeing and *epistemic* perception. The former kind of simple seeing is an ability that we share with many of our fellow creatures, whereas the latter is a capacity we possess, possibly together with some higher animals.[7] The difference between the two kinds of seeing is reflected in the way philosophers differentiate between *seeing x* and *seeing that x* is *F*. Suppose a man, who has never heard of the Western civilization, is shown a cell phone for the first time in his life. From his reaction we can tell that he can see the apparatus, he can touch it, circle around it, and even talk about it to his kinsmen, but he does not see that it is a cell phone. He can distinguish it from other items he is aware of, but this ability does not provide him with the ability to recognize it as a telephone. Without any concept of a cell phone his perception cannot supply him with the belief that he sees a cell phone. On the other hand, we immediately see a cell phone; we convert the information into a belief of what it is, as soon as we become aware of it.[8]

Quite analogous, a young student who is confronted with an X-ray picture of a body part may not see the black spot as a deathly tumor, while the trained physician, given his background knowledge, immediately sees it as such. But, as van Fraassen points out, it is silly to claim that we and the New Guinean do not look at the same object or that the student and the physician do not perceive the same spot.[9] Likewise, it would be absurd of me to deny that the spider and I are capable of seeing the same object

[7] Cf. Dretske (1963) Chap. 1, and his papers "Simple Seeing" and "Differences That Make No Difference" in Dretske (2000), pp. 97–112 and138–157. See also Brown (1987) Chap. 4.

[8] If I speak naturally, not in a philosophical setting, I would just say "I see a cell phone." To say "I see it as a cell phone" or to say "I see that it is a cell phone" seems to me to suggest the two-level sense data plus interpretation model of perception. This is certainly not Dretske's intention. Suppose someone makes a cell phone that looks like a wristwatch. I might, being presented with such a device, inspect it to discover what it really is and proclaim: "I see that it is a cell phone." But I would assert this latter sentence, rather than simply "I see a cell phone," only in the special circumstances where it is not, as it were, immediately apparent to me what it was that I saw. The two sentences are subtly different in meaning.

[9] van Fraassen (1980), p. 15.

when I perceive that it catches a fly. Both humans and animals are engaged in non-epistemic, sensory seeing, and the only requirement for such an engagement is that the organism has well-functioning visual senses.

Simple seeing gives us beliefs about color, shape, size, position, and movement, but not beliefs about the identity of things or the type of things. Dretske claims that animals and small children can see things, say a thimble, without having to judge them to be thimbles, identifying them as thimbles, or regarding them as thimbles. This mental situation consisting of simple seeing differs from the cognitive situations where we see that the chair is red, when the chair is in the room, whether it is in pieces, who sits on it, and what happens when the person rocks on it.

Perhaps infants and animals come sort of close to simple seeing—though some concepts may be so innate that they are imposed on the sensory given literally from birth. But I think once we attain self-consciousness our basic conceptual apparatus is so entrenched that "simple seeing" without engaging the higher perceptual faculty is normally not possible. So granting Dretske the existence of a mental state like simple seeing, and I do not see why we should not, seeing *x* still seems to involve some sort of beliefs, and therefore concepts, although perhaps of a very rudimental sort. Most probably we would not say that a spider has a concept of a fly, i.e., that *x* is *F*, and therefore does not convert the sensory information about its presence into a belief about the occurrence of a fly. Instead seeing *x* involves converting information about the presence of something, which in some higher animals can be described as converting information into a belief that *x* is *F*, where *F* refers to the property of 'being present.' Probably the evolution of spiders has not even brought it to this stage. Likewise, the spider receives information that something of a certain size and of a certain form is moving around. It seems reasonable to think that the information is then being compared to some genetically transmitted blueprint of edibles. If the size and the shape are 'right' according to the blueprint, which has evolved through selection and adaptation, the spider will attempt to catch the fly. Of course, somebody might say that what we mean by a "concept" is just such a mental blueprint, whether it is transmitted genetically or culturally. All I have done here is to say the same thing as 'concept' using the metaphor of a 'blueprint.' However, I do not accept such a reproof since the argument is actually meant to show that what we take to be concepts have evolved out of such neural blueprints. The crux of the matter is rather at which stage of evolution such blueprints went from being purely 'neural' to be 'mental.'

Consequently, if this is true, there cannot be a difference between non-epistemic seeing and epistemic perception that makes an epistemic difference. The role played by the epistemic component is a matter of degree. It seems to be an important fact about *epistemic* perception that most of what we see is directly recognizable as being this or that item. Walking into the wood I see trees, bushes, plants, leaves, grass, soil, animals, and birds as what I am immediately seeing. In case I focus on a tree, I immediately have a non-inferential belief that what I am looking at is a tree. I am not in a situation in which I have to interpret phenomenal impressions of green and brown colors to see it as a tree. I do not analyze the various phenomenal features of my perception in order to get the idea that the object I face right now is a tree. Seeing a tree is one and the same as gaining an immediate belief that there is a tree in front of me. However, the non-inferential character of this belief is dependent on my level of conceptual sophistication and the perceptual context in which I experience the item in question. I may directly see the tree as an oak, if I can distinguish between oaks, beeches, and other kinds of trees, or I may merely see it as a tree if all trees look alike to me.

Consequently, it seems appropriate to define epistemic perception in terms of belief acquisition: For all *S*, *x*: *S* perceives *x* as *F* if, and only if, *S* acquires a belief through information from the senses that *x* is *F*, where *S* is any subject, *x* is any object (or property) and '*F*' is a genetically or culturally transmitted concept or mental blueprint. If the subject does not gain such a belief, it means that the subject cannot apply the proper concept '*F*' to what it is non-epistemically seeing, and thus it will be epistemically unable to perceive the object. Perceiving is pre-linguistic—the dog recognizes his master—a single indivisible sensory-cum-mental event. Higher animals, no doubt, are capable of recognizing and identifying parts of what they see, for instance, a dog recognizes strangers, whereas people also seem normally able to know that they are seeing what they see. They may know that *x* is an instance of '*F*,' or they may be ignorant of what *x* might be because they lack a concept of *F*. As long as we move around in an environment in which we can put names on objects and properties we are seeing, we are involved in a more advanced form of epistemic perception. I may, indeed, know the names of different species of trees without knowing the circumstances under which they apply. In such a situation I do not (and cannot) understand that the tree in front of me is an oak. But I am still involved in epistemic perception, since at least I see a tree and am aware of seeing one. Even the aboriginal from New Guinea—who recognizes immediately a solid thing he cannot identify—is to a certain extent involved in a kind of epistemic perception by knowing the surroundings of the cell phone, its colors and form. He also recognizes that it is not a plant or an animal, that it is one object, that it is different from other objects, and that it is, in general, not an instance of any of the concepts that he does have. He has, as it were, categorized it into the category of "unknown things."

We may learn some facts supporting an epistemological analysis along these lines from patients with split-brain surgery.[10] One fact is that the linguistic center for most people seems to be mainly located in the left hemisphere, whereas spatial orientation, musical appreciation, and visual pattern recognition are located primarily in the right hemisphere. For example, when objects are presented to the left visual field of a patient with such split-brain surgery, all of the information about the objects is transmitted to the 'non-verbal' right side of the brain with the result that the patient cannot name the objects, even though the patient can recognize the objects by correctly picking them out when so directed. The person has some beliefs but not the linguistic capacity to name the objects that those beliefs are about. Thus, there is some evidence that the different functions of the two hemispheres may be characterized as two forms of epistemic perception, one in which objects are just recognized and another where the object is known to be what it is seen to be.

There is, in other words, no logical connection between our competence as a speaker in a certain area and our ability to see something as something specific, i.e., as an instance falling under a certain "concept" (or fitting a mental blueprint). The connection is close, nevertheless; without the conceptual sophistication, which our linguistic competence both presupposes and promotes, we would be incapable of seeing most items

[10] Cf. Sperry (1970). See also Marks (1980).

as something that belongs to a certain category. Some, more experiential concepts, which were relevant for our experiential orientation in our predecessors' environment, are innate because of adaptation.[11] Also infants learn the linguistic part by imitation, not by instructions, but that is only after the perceptual machinery has been engaged to the extent that it is able to identify different objects as falling in different categories, and all the cognitive apparatus that presupposes. However, we learn more advanced concepts by linguistic instructions about using the corresponding name. In these cases it would generally be odd to argue that a person has a belief that what he is seeing is, say, a wedding, but he does not know under which circumstances he may correctly name an item in his visual field a "wedding." Similarly, it would be incoherent to claim that he does know these circumstances and believes that he is seeing a wedding, but that he does not realize that the appropriate circumstances for applying the term are fulfilled. It would be equally absurd to say that he knows the circumstances that have to be fulfilled to use the term to refer correctly to an item, but that his perception of these circumstances does not lead him to any beliefs about what it is that he is seeing.

Indeed, it is obvious that beliefs we immediately acquire from our senses may be wrong, because the sensory information can lead to a mistaken belief. The conceptual converter may not be sophisticated enough. For years I see ravens every time I walk in the woods, but then one day an ornithologist tells me that the birds I have been seeing were in fact black crows. Since my linguistic competence was not developed well enough, I used the name "raven" incorrectly, and because of my mistaken concept of ravens, I directly saw ravens, where to be correct I should have seen crows. But most often people are correct in the beliefs they gain by the senses because, as competent concept-holders, they are able to apply successfully most of the names and predicates concerning their daily environment. If we were led to mistaken beliefs as frequently as correct ones, then our senses would be highly unreliable, and presumably evolution

[11] In some animals juveniles seem to learn concepts from experiencing their own species' behavioral actions (prairie dogs). We may even assume that before *Homo sapiens* developed the capacity of a language they had developed a rather sophisticated conceptual system in relation to their environment that was partly genetically inherent and partly taught from one generation to the next by sounds, signals, and actions.

would have filtered out those faculties. So while of course it is possible to be mistaken about what it is you are seeing, evolutionary theory supports the conclusion that this happens in only a minority of common cases.

3.2 From Perception to Observation

Evolution determines which things animals can sense and perceive as well as how they sense and perceive them. Humans are no exception. But the adaptation of our senses to our environment is harmonized with bodily movements. So the beliefs we gain from perception exist already in the brain as conceptual dispositions and were initially formed as a typological recognition on the basis of the coordination of various sensory impressions and behavior. Plants do not have sense organs since they are stationary. They get their nutrition directly from where they grow, and many use the wind, birds, or insects for pollination. In contrast, birds and animals move around to find food and meet a mating partner. Their senses guide their behavior, and their behavior guides their senses. Hence sensation and behavior have co-evolved in birds and animals. The brain is adapted to coordinate sensory information and proprioceptive and kinesthetic information. This is probably the reason why much of our sensory information is not stored as representational knowledge but is directly expressed in bodily responses. This saves costly reaction time, and the brain is not overloaded with information, which it does not have the capacity to handle, as it would be if both imagination and reflection were always required to behave efficiently. This also explains why embodiment is a key concept of cognitive theory. Embodiment refers to the fact that cognition results from not only processing sensory information in our brains, but also that our bodily actions can affect this information processing process.[12] Hence perception engages a disposition in which

[12] Bourdieu ([1997]2000) aptly describes the situation: "The world is comprehensible, immediately endowed with meaning, because from the very beginning, the body has been exposed to its regularities. Having thereby acquired a system of dispositions that are coherent with these regularities, the body finds itself predisposed and ready to anticipate them practically through behavior that activates a type of knowledge through the body, which ensures practical knowledge of the world… In other words, if the agent has an immediate understanding of the familiar world, this depends on the fact that cognitive structures, activated by the agent himself, are the product of an incorporation

the brain matches the sensory input with its repertoire of conceptual schemata for applying concepts (or mental blueprints), and when such a match is found it engenders a belief. A perception in itself necessarily involves an epistemic judgment or reflective awareness such as that I am *now* watching a tiger, a toilet, or a toothpick.

Thus, sensation and perception operate on two different levels. First, sensation works on the basis of certain innate dispositions to structure and enhance information coming to the sense organs. This is simple seeing. Some of this information may be blended with embodied information, i.e., information about the world that is kept as proprioceptive experiences, as when you drive a car through heavy traffic; your perception and bodily movement become one cognitive process. Some other is stored in our memory as propositional knowledge we can recall when we are in need. So whenever we see something as a fact, for example that the traffic light turns red, we are involved in epistemic perception. Such a fact concerns medium-sized objects that reflect the light from the sun or artificial light. But this is also where the range of human perception stops. Nature has not given us the capacity of sensing very minute things like atoms and molecules or even microbes or human cells. Neither do we have the capacity to experience things on a very large scale, nor do we have developed organs that can register electromagnetic radiation other than in the range of visible light. For instance, there is infrared radiation all over, but we have not developed senses for it, even though some animals have. Just because we do not have a receptor for some sensory input that does not imply that the sources of such input are few or that it is of no adaptive use. Of course, we have not developed X-ray or gamma ray receptors exactly because there is so little of that kind of radiation here on earth. However, thanks to human ingenuity we have been able to build instruments that partially compensate for our limitations. These instruments function just like natural sense organs. Some signals go into them in one end; humans receive information in the other end.

So we construct instruments to deliver, in a sensory form, information about something we cannot gain information about by unaided human

of the structures of the world in which he acts and on the fact that the instruments used to know the world are constructed by and through the world" (pp. 135–136).

sense organs. Collecting and recording this information give rise to beliefs about the object the information is about. Such a belief can be just that an unidentified *x* is present. In this case we want to interpret what *x* is other that an *x*. In other cases the belief we acquire is a belief that *x* is *F*. In these situations it seems correct to say that we are not involved in an act of interpretation.[13] *Observation* differs from *perception* in the sense that one must be accountable for those second-order beliefs one attains from observation by being able to point to relevant evidence in order for these beliefs to count as knowledge.[14] However, perceptual knowledge demands no such justification; the acquired belief only has to be reached by a direct sensory mechanism and to be logically compatible with what we already know. It is only if the reliability of a particular perception is put into question that we may find it necessary to defend it with further evidence.

At the time of the invention of the telescope and the microscope science moved from observation to instrumental observation. We still 'see' through one of these instruments because they use light to collect information, but we also observe stars or bacteria with them. Observation takes place when one concentrates on the object of one's perception or is epistemically aware of the information one receives from the environment. Therefore I shall propose an account of observation, which could be called *the internalistic theory of observation in contrast to the externalistic theory of perception*. Central to my account is the idea that just like perception observation involves belief acquisition. But where perception belongs to the embodied part of cognition, observation is due to the reflective part. Or to put it in a well-known slogan: perception does not require that you know that you perceive, whereas observation requires that you know that (and what) you perceive. If we grant that some dumb animals have beliefs about their environments, they may be involved in perception but they cannot intentionally contemplate the belief they received

[13] Faye (2012), pp. 98 ff.

[14] It is not obvious that the evidence has to be "adequate" but it must be relevant. It seems to me that observation can and sometimes does lead to beliefs that are mistaken. When Herschel first observed what we now call Uranus, he interpreted it to be a comet. Of course, he was making observations. So we might make an observation and might be able to point to evidence we might think is adequate to justify the beliefs we get from this observation, but that in fact this evidence is not adequate, and all we have is a mistaken belief. In other words, our mistake was about the adequacy of the evidence. But this does not mean that an observation was not made.

from perception. Animals cannot make reflective judgments about their beliefs. In science, however, a purely epistemic theory of perception is insufficient to explain *observation*, much less *instrumental observation*, and must be supplemented by a reflective account of observation. My contention is that the belief acquisition we find in instrumental observation can be grasped along the same lines as what we find in perception. The generalization will work if we realize that the common epistemic element in both perception and scientific observation is the acquisition of beliefs regarding the object the information is about. They differ only with respect to the degree of epistemic justification.

3.3 Theory-Ladenness

It is part of the folklore of today's philosophy of knowledge that observation is theory-laden. The phrase "theory-laden" leaves more room for culture and less for nature. Nevertheless, the expression is an unhappy idiom. It seems to ascribe more cognitive power to interpretation in terms of a scientific theory held at a given time than to the innate mechanism underlying observation. Nevertheless, when Norwood Russell Hanson introduced the term in *Patterns of Discovery* he was careful to distinguish between what he called "theory-ladeness" and what he called "the interpretation thesis." An interpretation is something put onto a supposedly given sensory input, something that in principle could be stripped away and leave an "uninterpreted, pure sensory given." I think with "theory-laden" Hanson had in mind something more akin to a Kant-like pre-conscious synthesis that was *inseparable* from what it is that we directly experience.

Already from Gestalt psychology it is known that the brain adds in its own information to the sensory input. The stimulus that arrives in the brain from one's eyes is the same regardless of whether one sees a white vase or two black opposing faces (Rubin's vase). Apparently, the brain processes this sensory information by finding a match between it and information of previously established cognitive blueprints, which in this case can be either a vase or two faces.[15] It does so automatically even though we might have established reflective knowledge that contradicts that perception. Think merely of the stick submerged in water that is seen as if it is bent, though known to be straight. Faced with a Müller-Lyrrel diagram we know by measurement that both lines are the same length. Nevertheless, such positive knowledge is not capable of totally removing the perceptual illusion of seeing one of the lines as longer than the other. A classical series of experiments made by the Belgian psychologist Albert Michotte on visual perception of causality demonstrates that in some situations people can have experiences of moving objects that are quite paradoxical when the movement is described according to ordinary mechanical principles.[16] In other experiments, made by two Danish psychologists, Edgar Rubin and Edgar Tranekjær Rasmussen, where the subjects are introduced to some simple geometrical figures, they feel constrained to give a perceptual description of the relationship of these objects although they are aware of it as inconsistent with Euclidean geometry.[17] Thus, established reflective knowledge does not override our perception in every situation. There are cases where relevant knowledge does not enforce itself on the perceptual organization even though it might form a significant part of our general background knowledge. Eventually we may learn to "correct" these illusions, say by knowing how an unbroken oar in water looks like.

The above examples illustrate that sometimes perception and reflectively pre-established knowledge are not consistent with each other. Background knowledge cannot always affect the way the brain processes the sensory information, which is a natural capacity of living beings and which may change from species to species. Reflective knowledge may merely affect, or rather constitute, the *observational* component, in addition to the perceptual component, because observation needs to be not only biologically reliable but also epistemically justifiable. So obviously there must be odd cases where these two kinds of elements are not, or cannot be, brought into mutual consistency. What is correct about claiming that observation

[15] Probably this is also the mechanism behind association and the Rorschach test. The sensory information is the same for every test person but various people process this information in virtue of different conceptual modules determined by their psychological character.

[16] Michotte ([1945]1962), p. 87 and p. 125.

[17] See Rubin (1950), and Tranekjær Rasmussen (1955).

is theory-laden, I believe, is that our reflectively established knowledge may have a direct formative influence on what we perceive something to be. It is a trivial fact that seeing an object as an object or as a certain object, say, a cat, requires a certain amount of conceptual resources. Generally, we never see something unspecific as "sense data" or sense impression, which then is interpreted in accord with whatever concepts are available to that perceiver. Although psychological research has shown that people never actually experience some kind of 'raw' purely sensory input, it is mistaken to conclude from this fact that all observation is laden by the scientists' theoretical commitments. While it is no doubt true that scientists strongly committed to a theory will tend to make observations directed by that theory, it is surely mistaken to assume that no scientist can ever be sufficiently open minded about his or her theoretical commitments to be able to describe perception from the point of view of common sense. Thus, it goes much too far when Hanson claims that our concept of seeing not only implies that we perceive something as this or that, but also that such or such *would be* the case, if what we do actually see were to be in a certain counterfactual situation.[18] As he says, "to see an X-ray tube is at least to see that, were it dropped on stone, it would smash." Are we really able to see counterfactual situations? Is it not so that the smash of the X-ray tube, if, contrary to the facts, it were to hit stone, is something we know, assume, imagine, or expect? Of course it is. We do not see it because this is not that about which our sense organs actually inform us.

As noted earlier, Wilson distinguishes between primary and secondary "epigenetic rules" corresponding to a distinction between the capacity of sensation and that of cognition. The function of sensation is to collect information and to distribute this information to the cognitive centers of the brain, whereas the function of cognition is to use this information to imagine what is going to happen next and to steer the brain's motor-control system of our body accordingly. Sometimes sensation and cognition are regarded as parts of a single continuous information processing involving recognition, identification, and classification. But from an evolutionary point of view it seems reasonable to keep their functions

[18] Hanson (1958), pp. 19 ff.

distinct.[19] Therefore, the claim that perception and observation are theory-laden only implies that a cognitive process somehow transforms the sensory information under the effect of what we already believe so that we not only see something but also usually see it as something.

What is correct about claiming that observation is theory-laden is that observation depends partly on a person's conceptual knowledge and linguistic competence and partly on the physical circumstances in which the act of perception takes place. It is surely mistaken to hold that we first receive information and then make a theoretical interpretation of this information before we infer what it is that we really observe. The processing of the sensory information should not be seen as a process of 'interpretation,' which happens only in case we consciously acknowledge a lack of information. Most of what we observe instrumentally does not go through an act of interpretation before we are aware of what we are observing. Exactly the same holds for observation with unaided eyes. Human beings become adapted to recognizing differences and similarities in the sensory information input as signs of certain types. After we have learned to distinguish various types, we no longer need to interpret what we observe. We are no longer just seeing something, but seeing it as a certain type. Familiar things and events are usually identified immediately as certain things and events whenever they are perceived in the appropriate circumstances. Moreover, as we learn to master a particular language that provides a sophisticated conceptual system, we see these things and their properties in the categories of that language. Our natural language provides us with the names for the things and events found in our everyday surroundings. We are immediately *aware* of a certain entity as a person, a house, a tree, or a mountain whenever we grasp that such an entity belongs to the type that has that name in that particular language, for if one's language contains terms for these entities, the mastery of that language consists in one's cognitive skill in unhesitatingly applying each term to the right objects.

[19] Also Dretske defended such a distinction as one between non-epistemic simple seeing and epistemic perception. He then identified simple seeing with information encoded in an analog form and perception with information encoded in a digital form. See his paper "Sensation and Perception" in (1981), pp. 135–153.

3.4 Instrumental Observation

The items in our vision are perceived as definite objects of a certain type because we can grasp them conceptually as such. It also seems reasonable to assume that the higher animals have perceptual beliefs because up to a certain point they can identify particulars as definite particulars or particulars as instances of definite types. The meaning I have reserved for "observation" implies that no higher animals, except human beings, have the capacity for making observations, because they are not able to justify the beliefs they obtain through their senses. Beliefs we gain by instrumental observation are, of course, only possible in a more developed state of human culture. Instrumental observation requires that the sensory information not only produces first-order beliefs about what one perceives, but also second-order beliefs about how the first-order belief can be justified with reference to the technology employed in extending the limits of our senses. In science we rarely have the possibility of comparing recordings we can see by the unaided eye with the items themselves. Most stars and nearly all galaxies are such items; cells, molecules, atoms, hadrons, and leptons are others. We cannot check alleged recordings of these objects with the objects themselves. When we compare pictures of Barak Obama with the man himself, we do it by seeing whether the person on the pictures has the same features and properties as the real Obama. And if there is a high degree of similarity between the pictures on the TV screen and the man himself, the producer regards it as evidence that the broadcast is working. But with respect to distant celestial objects or small corpuscles, which are either too faint or too small to be seen with the naked eye, scientists cannot compare pictures taken through telescopes or microscopes with the perceived features of the objects themselves. The same holds, of course, for other instrumental recordings of these items. When a skilled physicist looks at a photograph taken of a track in a Wilson cloud chamber (today this instrument has been replaced by a bubble chamber or a spark chamber), she immediately acquires the first-order belief that she is seeing a picture of a cloud chamber track. She may also directly acquire a first-order belief that what she sees is a picture of a track of an electron in spite of the fact that she cannot see the object itself and thus cannot compare it with the photograph. As an observation, however, she would also have obtained a second-order belief about the evidence supporting her first-order belief. Such a second-order belief is an epistemic belief. The difference is that first-order beliefs are perceptual beliefs, whereas second-order beliefs are epistemic beliefs that can be used to justify first-order beliefs.

Nobody would say that the physicist perceives the electron. But he perceives its track. In the same way as a photograph of something is not identical with what it is a picture of, a track or a signal is not identical with what has caused it. The track on the photograph has properties that the actual track in the cloud chamber does not have, and the actual track has other properties different from those of the electron. The electron itself, as distinct from its track, does not have properties that are directly perceivable to the physicists or us. Yet, as soon as he looks at the photograph, the trained physicist may acquire the first-order perceptual belief that the white line is a track of an electron. The perceptual belief need not be a result of an inference or interpretation of other perceivable features; he may immediately perceive the track as one of an electron. In addition, the sensory input produced by the instrument is converted not only into a first-order belief of the physicist that he perceives white line patterns on a photograph, but also a first-order belief that he is observing tracks of an electron. So therefore the question we have to answer is how can the physicist acquire such immediate first-order beliefs about something that is invisible? He can do this because his second-order beliefs are about phenomena that he takes to be not only good evidence for justification of his first-order belief but also part of his concept of an electron that white lines on a cloud chamber photograph of this nature are good evidence of the presence of electrons.[20] The white patterns on the photograph are conceptually associated with the tracks of electrons due to the fact that part of the concept of an electron is that it has properties that allow it to interact with water molecules in a supersaturated water vapor leaving behind a trail. Hence, the physicist perceives the patterns as tracks of electrons. The second-order belief is about the perceptual criteria that the physicist may use in this context to identify the bearer of the name. When the physicist sees that these perceptual criteria are realized in a certain experimental context, he or she can justify the belief that these patterns are made by electrons.

[20] Cf. Faye (2002) p. 38 and p. 74.

On the basis of her general training in particle physics and experience of the physical context in which such lines appear on the photograph, a skilled particle physicist will immediately perceive them as tracks of a certain particle. Her knowledge about the technical design and construction of a Wilson cloud chamber, knowledge of the experimental conditions under which the lines are recorded, knowledge of atomic physics, and in particular knowledge of how different particles leave different track patterns on the photograph, all this information enters the processing of the information she actually receives through her senses, providing her with the first-order belief that the line she sees is the track of an electron. All her knowledge enters into generating a second-order belief on the basis of the incoming sensory impressions and the first-order perceptual belief, because the physicist's notion of an electron includes the patterns on a cloud chamber photograph as one type of evidential criteria by which he can justify believing that an electron passed through the chamber. Given his background knowledge and knowledge of the circumstances under which this particular experience appears, the physicist is capable of both assigning the name of the particle that made the track correctly and using it as evidence to defend his perceptual belief, if needed. Since the young student does not know under which circumstances she may apply correctly the various names of fundamental particles, she is not able to form any perceptual belief about what it is that she is seeing, and she is not able to recognize the pattern on the photograph as a picture of the track of some specific particle. The student may even possess advanced knowledge of atomic particles but still have no knowledge of the experimental circumstances under which she can refer to an observation or a detection of any of these particles. Without this knowledge she is not able to see the pattern on the photograph as a track of a certain particle.

Here I want to emphasize that whenever the physicist can use a term like "electron" to refer to the perceivable outcome of an instrument as the effect of what the term refers to, he is involved in an act of observation of the referent of that term. As a competent speaker of the physicists' language, he knows the particular experimental circumstances in which "electron tracks" applies correctly to the lines he can see, and as soon as he realizes that he is confronted with these circumstances, he gains a second-order belief about the cause of information of his first-order belief. The knowledge needed to have such a competence comes from many sources. Many different experiments have previously identified the patterns created by the various kinds of particles. Theoretical knowledge such as the Lorentz force law has been used to make these identifications. Eventually practical knowledge gained by actions, manipulation, and experience together with theoretical knowledge of particle physics enables the physicist to become better and better at identifying the various tracks. Finally, when all this information is internalized in the physicist's mind, her observational competence has reached a conceptual level where she can apply the name "electron" correctly to what she sees whenever she is in the perceptually correct circumstances. However, she would not be able to observe the electron, unless she could actively use this background knowledge to justify her belief.

Since scientists learn of the existence and properties of objects that cannot be perceived through instruments that make them perceivable, I regard it as reasonable to say that under such circumstances where a scientist forms a second-order belief with respect to the acquired perceptual belief, he or she is 'observing' what cannot be perceived by the naked eye. Thus, we always perceive items in a context, and phenomena are understood or explained by science only in such a context. The white lines on a photographic plate become understandable to the physicist as a picture of the track of an electron when he knows the conditions that have to be satisfied in order for him to refer to the conspicuous delineations as tracks produced by the electron, and he is aware that the context of the phenomenon corresponds with these conditions. I hold that this awareness is a case of observation where the second-order belief gained by the very skilled physicist is non-inferentially acquired. But the belief also may be inferred indirectly from other beliefs because the observing person has too little experience, or because the item of observation has not been observed before, or because the reliability of the instrument is unknown. Whether or not the student perceives a white line on a photograph as a track of an electron depends on whether or not he knows about the ionization that happens along the path of a charged particle passing through oversaturated vapor in a container. If he does not know this, and does not know how a cloud chamber works, he will not know that the term "electron" is applicable in the given context, and he will not obtain the same perceptual belief as the skilled physicists. Thus, he will not acquire a correct belief about what he is seeing, and hence he will not epistemically observe an electron.

Indeed, a student or technician could be trained to identify the particles causing different tracks just on the basis of their shape and never know anything about the actual physics that causes the track to be generated in the cloud chamber and the ways in which differently charged particles behave under different forces. In this case the student or the technician is not observing anything. This identification could even be carried out by a computer. If I am correct, the student or the technician lacks relevant secondary beliefs to justify their first-order beliefs. When stellar spectra were first measured astronomers had no idea what caused them, but they could still say they had "observed" certain elements in the sun's atmosphere simply because they could match the lines with known samples of the elements.

Instrumental observation is *an intentional act* of getting perceptual information about an object in order to find out whether the perceptual conditions for the application of a certain name to the phenomenon in our visual field are satisfied. These perceptual conditions concern the evidential criteria for applying that name. Therefore, in contrast to my definition of perception, I will say that we are making an 'observation' if, and only if, on the basis of perception, we come to believe that the conditions for assigning a name or a predicate to objects that we cannot see have been met because those criteria we can see are fulfilled. In other words, for all *S*, *O*: *S* observes *O* as *N* if, and only if, *S* forms a second-order belief that the object of his or her perception satisfied the evidential criteria for justifying the use of the name "*N*" and *S* is justified in having this belief.

As a competent speaker of an explicitly defined professional language, the scientist knows the circumstances in which the names and predicates of that language apply correctly to something in his visual field. But the scientist would never have become a competent speaker if she could not recognize those circumstances whenever they were present. However, recognizing that one is in the appropriate circumstances that determine the correct use of a certain declarative sentence—a sentence containing some of these names and predicates—is the same thing as acquiring a *warranted* belief that the sentence can be asserted.

A little more can be said about the circumstances under which the correct application of observational terms becomes established. The sense of scientific terms is specified by the coherent structure in which the terms are first learned and then employed. I hold that various observational and experimental practices fix the reference of observational terms. These practices both identify the bearer of the names and determine whether the conditions are satisfied for assigning a certain predicate to a particular subject, which bears those properties. Thus, a term can function as an observational term, independently of whether or not it refers to something visible to the human eye, if its application can be associated with certain canonical experiments and observational situations. I do not hold that the meaning of such a term is identical with certain instrumental procedures as the operationalists once claimed. But the existence of canonical procedures makes certain that we can establish the referent of the terms referring to invisible entities, and by doing so we are genuinely observing these entities. Recognizing these procedures is what warrants a belief that we can observe what is imperceptible and gives us the warrant for making true statements about them.

3.5 Observability

The nature of observation consists in forming an epistemic belief concerning the perceptual items we are confronted with in our visual field and, if necessary, the items that are the objects of our instrumental recordings. According to this characterization of observation, every item that is able to cause such an epistemic belief is an observable object or property of an object and every item that is unable to do so is unobservable.

In defending constructive empiricism, van Fraassen puts a heavy burden on his distinction between observables and unobservables. Rightly so I would say, but he puts it up in the wrong place. He maintains that the alleged distinction is in no way a clear-cut distinction and that many borderline cases exist between what is observable and what is unobservable. In effect, he grants that "observable" is a vague predicate. However, he also holds that there are clear cases and counter cases for determining its use.[21] Seeing with the naked eye is the paradigmatic example of observation.

[21] van Fraassen (1980), pp. 16–17.

Also a look through a telescope at a moon of Jupiter is clearly a case of observation, because it is possible for astronauts to go out there in a space shuttle and see the same thing for themselves with their unaided eyes. The counterexample is the purported observation of an atomic object in a cloud chamber. In this case, van Fraassen holds, a particle is detected by means of the cloud chamber, if the current theory is correct, and then this detection is referred to as an "observation of an electron" by the physicists. What they see is the track caused by the particle similar to the vapor trail in the sky left by a jet. The particle itself is unobservable. The same is true for things that can only be seen through a microscope.[22] Such microscopic objects are unobservable, according to van Fraassen's distinction, contrary to telescopic objects a great distance from us, since they can never be scrutinized directly by human eyeballs.

For van Fraassen, the distinction between what is observable and what is not is fundamental to empiricism. However, he insists that the dichotomy has no ontological implications, because observability is too anthropocentric a notion to entail anything about what is real and what is not. But it has a very important epistemological role to play by telling us that belief in truth is restricted to observability.[23] Nevertheless, van Fraassen seems to ignore that to say that an object is observed is to say that certain beliefs about unperceivable objects are true, and if these beliefs are true, then the referents of the statements must be real objects. However, van Fraassen concludes, we humans should not believe a theory about something that we cannot observe to be true. We are only justified in believing the theory to be 'empirically adequate' in case what it says about what we can observe is true. So accepting a theory means that one is willing to claim that it is empirically adequate but not necessarily that it is true.

Obviously, van Fraassen's argument for empiricism contains two independent elements: (i) observability is coextensive with a belief in truth (not truth itself), and (ii) something is observable if it can be seen by humans, that is, it is observable if there are circumstances under which that something can be visually present to us. The second assumption is distinct from the first. A person may accept (i) while she denies (ii), or she may hold (ii) but not (i).

[22] See van Fraassen's answer to Ian Hacking in van Fraassen (1985), pp. 297–300.

[23] van Fraassen (1980), p. 19.

Here I shall not address the first assumption; it concerns a topic that I will discuss in a succeeding chapter. Here I want to raise questions about whether van Fraassen's characterization of observability is adequate.

Indeed, we cannot see electrons with our naked eyes simply because they do not possess properties that we have evolved to be able to perceive. All those properties that physicists attribute to electrons imply that they can be observed only by instruments, that is, electrons can be attributed only properties that are revealed in its causal interactions with macroscopic instruments that we can perceive. What we can see are traces (i.e., causal effects) of an electron or a record of these traces. Physically speaking, observational beliefs are acquired indirectly through the medium of 'observations' by means of instruments that extend human sensory limits. In this perspective instrumental observation is not so different from ordinary perception where visual beliefs are obtained only indirectly through the medium of the sunlight's reflection from the object.

I want to draw a distinction among (1) things that can be seen with the naked eye, i.e., things that are *perceivable* (or *simply observable*, if they give rise to a second-order belief), (2) things that are *prosthetically perceivable* (or *prosthetically observable*, if they give rise to a second-order belief), and (3) things that are observable with respect to the art of instrument making, which may be called *instrumentally observable*. The first is more or less biologically constant as long we are restricting ourselves to humans, although also concept-dependent. The second and the third kinds of observable things depend on our level of technological development at a certain time.

There are two senses of observe, *synchronic* and *diachronic*, one might call them, and therefore also two senses of being perceivable or simply observable. Sometimes—when we are being very strictly literal—what we 'observe' is taken to refer to 'at an instant.' As it were, a frame in the running movie of time. Here change is inferred from a comparison of many observations over time. However, when we observe change itself, dynamic processes with temporal thickness, then of course we add together many frames of the movie. In this context we observe 'processes' more directly than the momentary properties of objects, as we might "observe" the rebuilding of a city after a bombardment or one run of an experiment consisting of thousands of runs.

Things that are prosthetically perceivable are those that can be experienced by a person only because a technical device is implanted into the body in replacement of a sensory impairment. In such a sensory substitution information from an artificial receptor is coupled to the brain via a human-machine interface. The brain is then able to use this information in place of that normally transmitted by an intact sense organ. What separates prosthetic perception and instrumental observation is that the first form is normally not an intentional act; usually it produces only a first-order belief, in contrast to instrumental observation, which in my opinion always involves a second-order belief. But in cases where prosthetic perception produces a second-order belief by which a person can justify his or her first-order belief, prosthetic perception becomes an instance of an intentional act of observation.

Now, imagine some astronauts who travel out to Jupiter and take a look at one of its moons. Their perceptions at close range can confirm observations made by an earth-bound telescope. According to van Fraassen, this means that humans see through a telescope whenever we could bring ourselves into situations in which what is seen through that instrument will be immediately present to us. Thus, van Fraassen subscribes to (3) if, and only if, it is in principle accessible to (1). But how is "in principle" to be understood here? Since van Fraassen leaves it to science to tell us what is or is not observable, he cannot mean logically possible.[24] It seems therefore that van Fraassen holds that a thing is observable if, and only if, it is perceivable by the naked eye because that thing physiologically can become an object of simple seeing.

Nevertheless, not all events or things beyond our planet are physiologically possible for humans to encounter. Quasars billions of light years away or supernovas appearing in the Andromeda Galaxy may be so distant that they would have faded away long before we could arrive at their vicinity of space, so we could see them with our own eyes. This is not a question of whether or not it is within the range of our current physical technology. Some astronomical phenomena, like quasars, may belong to a certain category of objects that exists only in the early history of the universe. So even given the speed of light and time enough to travel out into the farthest region of space, it would be physically impossible for humans ever to confront such object. We may begin our journey to get there only

[24] Ibid., p. 57.

to see a quite different kind of object than the one we observe in the night sky today when we get within a range visible to the human eye. So taking seriously van Fraassen's criterion of observability amounts to a situation where some of what we 'observe' by looking through a telescope, like galaxies, is observable but some other things we 'observe,' say, quasars, are not, in spite of the fact that these celestial things look more or less similar when observed through a telescope. By refusing to rely on information that only scientific instruments yield, van Fraassen supports a concept of observability according to which something is observable only if human beings have the capacity of simply seeing it. Being 'observable' becomes a highly species-specific relation. This seems to be a very Darwinian attitude, and so it is. But if this is the only thing one has to say about observability, one ignores that biological selection has given the human species the ability to learn from experience, among other things, to construct instruments for collecting empirically accessible information.

Van Fraassen is correct in insisting that the physicist does not see atomic particles in the cloud chamber, but she may see tracks of them. Similarly, whenever a person looks at some footprints of a bear in the snow or at a photograph of a very faint and distant galaxy, he does not see the object itself but sees a trace of (parts of) it. He experiences only traces left by the object whose existence he thereby regards as revealed by these traces. In both cases the object may have ceased to exist long before its traces have disappeared. The same is true for atomic particles passing through a cloud chamber. In his criticism of the causal theory of perception, Dretske points to a useful analogy.[25] The difficulty with this theory, says Dretske, is that we cannot identify which part (the differential conditions) in the causal chain that we actually see. His illustration is Herman ringing a doorbell. How do we explain, for example, that what is audible and what is not audible in this process? The causal connection is a necessary condition only. Dretske argues that we must add an information component to the causal process. What matters is that a causal sequence carries information. Different causal sequences can carry the same information about a remote causal factor without carrying information about a more immediate and intermediate causal factor. We hear the doorbell ring, but we hear (audibly

[25] Dretske (1981), Chap. 6.

observe) that someone is at the door. The reason is that we hear that of which our hearing gives us the primary representation. This primary representation of the properties of remote objects is dependent on the constancy with which they are represented. The same holds for all forms of perception. The physicist perceives the lines as the track of an electron, but he observes that an electron went through the cloud chamber, because if he is challenged he may point to the lines as directly visible evidence.

Sometimes we might see a visible object directly indicating the presence of some invisible object, as a jet's trace in the sky indicates the presence of a jet we can barely see. But other times we might see that same object directly as the object itself, as the jet trace, or as a long streak of ice crystals in the stratosphere, which might interest the meteorologist. Atomic objects, viruses, and bacteria may therefore be observable in the instrumental sense, since there are experimental situations where they are represented by their effects on a screen or a display. However, one may object that we can always distinguish between things perceptually, as we see them, and the visual traces they leave behind them. But how can we distinguish between invisible entities and their various perceptual traces, which are taken to provide us with information about them? The dilemma is this: the invisible entity is distinguishable from the visible effects that are considered as evidence for its existence or it is not. If it is, then we seem not to be other than inductively justified in asserting that these entities really exist, or we can make no such an epistemic distinction between invisible entities and their perceptual traces, but then it seems as if invisible objects are nothing but these effects.

Ignoring black holes and perhaps a few other similarly problematic objects, macroscopic objects are indeed simply observable. We see these things because they have certain properties that make them perceivable for us. At the macroscopic level the properties of size, shape, and color are certainly attributes that allow us to have perceptual access to those objects that possess them. The perceptual beliefs we acquire about macroscopic objects can be justified in terms of other beliefs. Nevertheless, our sensation of these properties is of little interest from a physical or chemical point of view. Size, shape, and color of objects can be, and certainly are, of interest to physicists and chemists. But it is their measurement that counts in science, so instruments are needed for measuring the values of these properties.

The scientist has to abstract from the immediately perceptual properties whenever she wishes to give a full physical description of the system that she can measure. In such cases she must assign to the system other observable attributes whose numerical values are not directly perceptually accessible in the sense that they can be picked up by the unaided eye. What we might pick out with the unaided eye is the comparison of some visual property of an object with some measurement scale in virtue of which that property is accorded a measured "value." For instance, the mass of the system can only be experientially apprehended by a lever, its charge can only be recognized by an electroscope, etc. Physically relevant properties that we can see directly with our own eyes are very few indeed. The position of macroscopic objects, or the change in their position, belongs to this category, but even position is very often determined by instrument readings. Aside from the measurement of spatial and temporal intervals, the great majority of all other attributes of any significance in physics and chemistry require the use of some kind of instrument to manifest any recordable trace, as when we read the position of the pointer of a measuring instrument, or electronic digits on a computer screen, or watch the graphic curves on an oscilloscope. Since nearly every observable attribute in terms of which science describes macroscopic objects in order to predict their kinematical or dynamical behavior is based on an instrument reading, it seems arbitrary to claim that these properties make the object observable only if they are attributed to macroscopic objects. The reason is that even though we also attribute microscopic objects the same sorts of kinematical and dynamical properties, such objects are not observable.

As we have already pointed out, the obvious counterargument to this conclusion would be to say that observables attributed to macroscopic objects by their corresponding pointer reading observables are directly perceptible themselves, whereas the so-called "observables" attributed to microscopic objects by their corresponding pointer reading observables are not truly observed. It is paradoxical, to say the least, that the first set of macroscopic observables has the status of genuinely observable quantities, whereas the second set of microscopic observables has the status of being only theoretical quantities, because the pointer reading observables are in both cases only those with which we can be visually acquainted. However, the counterargument loses much of its appeal when one bears in mind that macroscopic properties, which usually are considered as simple observables, can sometimes be assigned a numerical value in situations where they function solely as theoretical quantities, and not just as perceptually accessible pointer reading quantities.

First, we need to introduce a distinction between two different kinds of quantities, those that are correlated with a pointer reading observable and those that are not. In our terminology the first kind includes the instrumental observables, while the second kind includes the purely theoretical quantities. Now consider, for instance, a double-star system, or the Earth's movement around the sun. The mass and the velocity of these bodies cannot, not even in principle, be observed by the eyes of any human being. These attributes can be immediately experienced by our senses as long as we are handling medium-sized objects here on Earth, though their exact numerical value cannot be established without the use of instruments. But not even this will do in the case of a double-star system. Indeed, the relative position of the two stars can be observed through the telescope as well be perceived directly from a nearby location. However, the assignment of mass to the bodies is possible, not as something which can be perceived directly, but as a specific numerical value of theoretical quantities based on Kepler's third law in Newton's general formulation, $T^2(M_A+M_B)/a^3 = 4\pi^2/G$, where a is half the major axis in the relative orbit specified by the observation of their mutual positions over a period of time. Their velocity can then be calculated on the basis of the conservation of energy from the general formula $V^2 = G(M_A+M_B)(2/r - 1/a)$. The mass and the velocity of the stars, or the same properties of the Sun and the Earth, are strictly theoretical quantities, which can never correspond to some appropriate pointer reading observables. And the theoretical parameters a and r, representing the orbital characteristics, can be said to designate observable quantities only in a derivative sense, since in actual fact they are calculated on the basis of many position measurements. So what in one context presents itself as a simple observable property, because it can be assigned to an object by visual means, appears in another context as an entirely theoretical quantity. Apart from showing that observation is laden by theory, this example illustrates quite well that the distinction between observable and non-observable quantities is not absolute but varies relative to the limits of our perceptual and instrumental possibilities.

Nevertheless, this counterargument does not seriously threaten my claim that we do observe *imperceptible* entities because it entirely fails to make note of a distinction between my claim that instrumental observation necessitates acquiring beliefs about theoretical entities through our senses and instruments as well as the question of how one can be justified in believing what one observes. In the case of perception, we may still perceive whatever it is that we do perceive without having warrant for our beliefs. This holds for direct perception as well as for instrumental observation. All observation, including perception with the unaided eye, depends on background knowledge. What is essential to both perception and observation is that the subject acquires a belief based on sensory information. This information being shaped by a given conceptual blueprint determines the perceptual or observational content regardless of the purely phenomenological form in which it appears. The acquired belief may or may not be true, but truth value is not a property of some concomitant qualia. Hence a distinction between observable and unobservable entities based on whether or not an entity can be seen with the naked eye, and therefore has a phenomenological form, fails to have any epistemic significance. To attribute visual properties to an object of perception is nothing but to acquire the first-order beliefs that the object exists and has these properties. Similarly, an object is observable if it has some properties that make it observable; that is, an object is observable if we can form a second-order belief that what we see is good evidence for our first-order belief. So it is correct to say that scientists observe microscopic objects, like atoms and electrons, whenever they actually take them to exist by applying their names correctly to what they *see* visually or by attributing to them a certain property on the basis of the outcome of an experimental investigation. Any object or attribute becomes 'observable' for us when we, *Homo sapiens technologensis*, are able to construct instruments that provide us with information about their presence and character.

4

Theory and Reality

Since the heyday of logical positivism, the dominant view in philosophy of science has been scientific realism.[1] But over the last two or three decades its prominence has seemed to dwindle. No one wants to return to the excesses of logical positivism, but as the dust after the battle settled, it became increasingly clear that not everything the defeated party had defended was without merit. And, as we shall see, scientific realism has its excesses and problems too. Hard-core instrumentalists believed that scientific theories are merely tools for predictions and calculations and that they have no ontological content telling us how the world is, being solely conceptual tools that are neither true nor false. Theories help us to organize empirical data by postulating the existence of 'theoretical entities,' but theoretical entities are, and always will be, fictitious mental constructions, because their alleged existence would transcend anything that could be established by sense experience.

[1] Scientific realism was explicitly defended against antirealist opponents already by Galileo (Belarmino was clearly holding a kind of instrumentalism), and the consensus of the centuries that followed was that Galileo was right. So realism was already long entrenched before positivists began to doubt it. True, one could say the nineteenth century was a turn away from this traditional realism, but remember not everyone in the nineteenth century was an idealist.

J. Faye, *Experience and Beyond*,
DOI 10.1007/978-3-319-31077-0_4

Scientific realism grows out of common-sense realism and the practical and observational success of science itself. Instrumentalism and hard-nosed empiricism, in contrast, are generated by a philosophical desire to strip metaphysics of any veil of legitimacy and to dress science in the armor of epistemic warrant. As long as astronomy, physics, chemistry, and biology dealt mainly with macroscopic objects that could be observed, as was generally the case prior to the end of nineteenth century, the acceptance of an instrumentalist view had no far-reaching implications with respect to either the number of theoretical entities explained away or possible technological consequences of a belief in these entities.[2] But with the development of new theories about invisible entities, forces and processes such as electric and magnetic fields, molecules, and atoms, and together with the rapid increase in technology based on our beliefs in the reality of such entities and processes and the truth of what our theories tell us about them, it seems pointless to insist that we possess no knowledge of that part of reality that is not directly accessible to our naked eye.[3] It is, the realist would say, only because scientific theories provide us with knowledge of the hidden structure behind phenomena that we have been able to change nature, design new organisms, and improve the material and technological level of modern society. Science not only yields theories that predict how well-known phenomena may change. It also fosters theories that give us insight into the laws of nature—thus allowing the creation of quite new phenomena never seen before. Hilary Putnam once declared: Realism is the only philosophy that does not consider the empirical success of science a miracle.

Common-sense realism is our inheritance from our biological ancestors. Our cognitive apparatus has been adapted to think of the world

---

[2] Historically, this flies in the face of the usual opinions of the Galileo vs. The Church debate, in which Bellarmino defended an instrumentalist view of theories, specifically the Copernican view, while Galileo insisted on a robust realism, partially as a consequence of his telescopic discoveries. This did indeed have "far reaching consequences." I have no intention of denying the truth of this opinion.

[3] A great deal of nineteenth century scientific debate was also concerned with the "unobservable" reconstructions by geologists and palaeontologists of past ages. It is less tempting to take something like a tyrannosaur or an ice age as a purely instrumental construction designed to organize data. Here the—dare I say 'innate'—realist reflex is so natural as to make the antirealist view seem as far-fetched as it really is.

as separated from us. As long as we can push things around, create and destroy them, we are biologically determined to take our experience to be of things that are independently real. Even though human beings cannot see, say, atoms, we—living in cultures highly influenced by Western science—take them to be real as part of common-sense realism. Physicists can act upon them and their existence makes the outcome of human actions understandable. Biological adaptation has also arranged that we have the abilities to identify perceptions across temporal separations as different perceptions of one and the same object.

In the preceding chapter I argued against the naked-eye empiricism because it was not committed to the observation of invisible entities. However, here I shall take issue with some of the most common arguments in favor of scientific realism, which go far beyond what an evolutionary-based ontology can accept. My aim is to show that scientific realists who advocate *semantic realism* lack philosophical arguments in support of their *metaphysical* thesis that currently accepted theories must be true or approximately true of the world as it really is in order to be able to explain their empirical success. Similarly, an alternative form of scientific realism, called *structural realism,* which is very much in vogue these days, is not better off. Rather than being a realist concerning *theories*, the Darwinian shares company with those philosophers who are realists only concerning *entities*.

4.1 Forms of Realism

Any theory of reality is "metaphysical" whether or not it is alleged to be justified by rational or empirical means. Thus, materialism, physicalism, naturalism, phenomenalism, idealism, and solipsism are all metaphysical theories. The aim of any metaphysical theory is to give a true account of the ultimate nature of reality. Those metaphysical theories that take the nature of what exists to be independent of the mind are realist theories, while those that see the mind as having a constructive influence on the nature of existence are antirealist theories. What is characteristic about metaphysics is that it goes beyond science in the same way as science goes beyond common sense.

Metaphysics can be subdivided into *ontology*, the account of the beings or entities of which reality is composed, and *cosmology,* which is concerned with the structure and origin of those entities. Thus, whether one says "objects" are the ultimate beings or "structures" are ultimate, one is holding an ontological, and therefore metaphysical, view. Different metaphysical theories have claimed the primacy of one over the other, but it seems to me that it is difficult to talk about concrete structures without objects that enter into those structures or concrete objects without saying how they are structured with respect to one another. I would use "concrete beings" or "concrete objects" to refer to the particular beings of one's ontology; their opposite, "abstract entities," denote, inter alia, those "essences" or "universals" that can "exist in" or "participate" in many particular beings. In the Middle Ages, realists were those who maintained that such abstract objects could exist apart from particulars, whereas, of course, nominalists denied that. There could be, of course, 'particular structures,' such as the structure of the solar system, and 'abstract structures,' such as are studied in mathematics. But a structure remains 'abstract' until some particular objects enter into that structure. Physics could describe in detail an infinite number of abstract structures of possible planetary systems, but none would be 'concrete' unless some particular collections of stars and planets entered into them. We shall return to this issue at the end of this chapter.

All physical objects or objects claimed to exist in space and time are, however, concrete. (It would be, on this definition, possible to argue that "strings" or "other branches of a multiverse" are not physical objects, because they exist in other dimensions than the familiar ones of human experience.) Classical materialism, or mechanistic materialism, asserts that all concrete objects are ultimately and only physical objects possessing essentially and exclusively those properties that define the classical mechanical state. But the introduction of the concepts of fields and energy in physics eventually pulled out the rug from underneath classical materialism.

"Realism" designates a possible ontological position in many different fields. If one believes that the external world exists independently of our consciousness, regardless of whether or not one believes in the existence of particular things, one is a realist with respect to that surrounding reality. Or in case one is in favor of the claim that there are moral facts that are not, in some way or another, determined by people's sentiments and emotions, one is a realist with respect to the ontological status of moral values. Or if one takes the view that abstract entities such as numbers exist, even though they are not observationally accessible, one will be a realist concerning mathematical quantities. We can also be realists when it comes to kinds, universals, modalities, and possible worlds. Common to every realist concerning these different areas is that what he is a realist about is taken to be real, regardless of whether he himself or other human beings now or ever have existed. But it is not a requirement that if somebody is a realist in one area, she must be so in every other area. Thus, there is no implication between a belief in the independent existence of the external world and a belief in, say, the independent reality of abstract objects.

Here I shall make a distinction between metaphysical realism and scientific realism. *Metaphysical realism* claims that reality as a whole exists in itself independently of the human mind. Apart from concrete entities in space and time, it may also be made up of abstract entities such as universals, relations, modalities, and substances whose identity conditions are necessarily true. According to the metaphysical realist, there exists one and only correct way of representing the fundamental nature of reality, and it is the aim of this true metaphysical theory to give us a mind-independent description of what exists and how it exists.

In contrast, *scientific realism* holds that based on empirical evidence science is able to tell us something true about the world. One version, called *theory realism*, maintains that our best scientific theories give us a true or approximately true representation of the physical world in terms of some fundamental laws of nature and some entities governed by these laws. The aim of science is with the help of these theories to give us an explanation of what can be observed. The other version, *entity realism*, holds that science is able to discover the existence of invisible entities without being committed to a position that its theories are true or approximately true. The referents of these theories are models rather than real entities; thus entity realists can avoid the question of the truth of theoretical statements.

A third form of realism is *common-sense realism*. For the sake of my argument, at this point I want to make a distinction between *mind-dependence* and *species-dependence* that allows me to explain common-sense realism as a result of human evolution. Common-sense realism has it that the empirical world is real in the sense that things around us are not fictions of our personal mind. What we see or act upon continues to exist in situations where nobody is around. Because of all the traces found in our surroundings we are convinced that the planet existed before we were born and if we all died out at once because of a certain disease we are sure that mountains, rivers, and oceans would continue to exist. Nonetheless, since one may be a realist with regard to truth too, the obvious question is whether or not one can be a common-sense realist without being a metaphysical realist with respect to truth. Is it such that there is an ultimate reality, which makes our beliefs true independently of how human beings perceive the world? Before answering this question, we need to throw more light on the realist view that entities exist independently of our sensations of them.

Two currents flow through Western philosophy: rationalist and empiricist. There have been those rationalists who anchor their account of the nature of this reality in what they take to be the dictates of reason. And then there have been common-sense realists, Hume notwithstanding, who take their account of reality from the dictates of the senses. So I distinguish the rationalist and empiricist vindication of realism. The rationalist approach can be associated with various forms of Platonism, aprioricism, and transcendentalism, whereas the empiricist approach would most notably be represented by most varieties of materialism, particularly in the robust version allied with classical mechanistic atomism. But there are also, of course, other versions that hold diverse particular views about the nature of this external reality. Although both rationalists and empiricists hold that concrete entities exist in space and time, they are divided with respect to whether that reality includes more kinds of entities than concrete objects and processes in space and time. The rationalist-minded philosophers believe that reality consists of more than concrete objects and therefore are metaphysical realists with respect to both abstract and concrete objects; the empiricist-minded want to confine reality to only those things that exist in space and time. For that reason they are metaphysical antirealists with respect to abstract entities. Thus, *metaphysical realism* holds with respect to abstract objects such as numbers, universals, relations, propositions, or possible worlds that they are a part of reality and exist in their own right. However, *metaphysical antirealism*—as another ontological position within metaphysics—holds with respect to abstract objects that their existence is dependent on human beings. Antirealists, such as nominalists, reject the reality of abstract objects over and above concrete entities but they need not be idealists tout court. In fact, nominalists are metaphysical realists with respect to concrete entities.

The term "antirealism" is basically a newer addition to the vocabulary. Of course, nominalists are (perhaps the first) antirealists as the Medieval debate had it, but in more recent times the main form of antirealism has been the more full-blown idealism, that is those who maintain that the perceived object is wholly constituted by the subject. "Realism" in the twentieth century arose as a reaction against this view, and the issue came to be in philosophy of science the issue of whether or not the conclusions of natural science tell us anything about a mind-independent reality, or whether those conclusions are wholly a construction of the scientific community. In other words, *scientific antirealism* in the form of instrumentalism denies that the nature of an independent reality has anything to do with the conclusions of science or that we can make any inferences about it, even hypothetically, from science.

For scientific realists the nature of reality is not a closed issue as it is for the metaphysical dogmatists; instead it is a provisional hypothesis ever amenable to revision on the basis of scientific progress. So the scientific realists are bound only by the conviction that what we know about the nature of reality is founded solely on what accords with our best scientific conjectures confirmed by empirical evidence. However, theory realism is grounded in a representational view of scientific theories: For a theory to be true or approximately true, it must represent the world more or less as it is, and our best theory within a particular discipline *expresses* the knowledge we have gained about the entities covered by the discipline. Karl Popper was probably correct in asserting that no scientific theories, even taken collectively, can demonstrate the truth of any metaphysical claims, but they can falsify many metaphysical claims that have been made. I think this is the case with mechanistic materialism, once taken as having been demonstrated by classical physics, and the falsifying agent here is quantum mechanics. In this sort of view metaphysical 'truth' can be thought of those beliefs about reality that remain unfalsified if the critical process is carried to infinity. So I would consider it a "pragmatic" metaphysical attitude, open to future revision. Metaphysics must "accord" with our best physical theories, but cannot be proved by them. The nature of that "accord" is what must be worked out by those who defend a certain metaphysical view. It is a matter for careful philosophical diplomacy.

Let us begin with the realist claim of mind independence. The common-sense realist may have two different positions in mind concerning our ontological commitments. The first is that the empirical world exists more or less as we perceive it, which means that the world is as we see it independently of human subjectivity. The world is not constituted by our personal wishes or beliefs; whatever exists exists as we would see it regardless of whether or not you and I are aware of it. How else should we explain the discovery of new objects that no humans have thought of before their detection? Anything that is accessible to our senses may be real, even though it is not actually an object of our perception. The second position is that the external world contains more things than what can be perceived by our senses. The common-sense realist hereby makes the external world physical, or mostly physical. He may contend that the world consists of the familiar things that surround us in our natural life and of those things discovered by science. The physical world as we perceive it is a mind-independent world. The world consists of ordinary stuff and objects like gold, water, human beings, animals, cars, and refrigerators, but also fields, forces, molecules, and atoms. The first position could be called the *plain version* of common-sense realism. Also the common-sense realist may hold without any problem that the everyday view of reality has to be supplemented with the scientific story about invisible but "observable" things and their properties, a position that I shall called the *sophisticated version* of common-sense realism.

Indeed, a metaphysical realist is not prevented from holding that the mental is different from the physical, nor therefore from claiming that the mental is objectively real, independently of whether someone believes it or not. But this is not what common sense tells us; this is due to metaphysical reflections about the constitution of the world. Realism, as a metaphysical position, does not rule out objects like minds, but it claims that the existence of minds and their specific nature are what they are regardless of the way one actually may conceive and apprehend them and regardless of whether they are objects of anybody's apprehension. Thomas Nagel, for instance, believes that there are subjective facts that are unattainable to human knowledge.[4] The requirement of logical independence of human knowledge also means that things, events, and laws can exist even if they cannot be observed, that is, even if they are, in principle, empirically inaccessible. The realist must insist upon the possible reality of such entities. The world may be inconceivable to our mind. Nothing in his metaphysical point of departure excludes the existence of unknowable entities as a genuine possibility.

This brings us to the realist's thesis about the question of existence. Is a realist committed to believe in an ultimate reality? The *metaphysical* realist takes a big step further and goes beyond the external world of common-sense but also beyond that of scientific realism. He adopts a God's eye view of objectivity. In Kantian terms this requires that true metaphysical theories are able to describe real entities as they are in themselves and not only as they appear to us. However, the objection against this form of objectivism from common-sense realism is that we are unable to step out of our epistemic situation and describe reality independently of our own cognitive resources. Of course, anything like this would be impossible. We are part of the world. We cannot describe the world from the point of view of "nowhere." It must always be done from "somewhere." As part of reality, we must describe reality as our biology has adapted us to understand it. So the evolutionary naturalist holds that the nature of the external world is mind-independent but not species-independent.

The metaphysical as well as the scientific realist assume that the external world exists apart from any facts about human epistemic states; both deny that our thinking gives things their being. Knowing does not create being, but being makes knowing possible. So far the evolutionary naturalist agrees. But the metaphysical realist argues further that this implies that the world is what it is in itself, i.e., it has a *nature* that is what it is independently of any cognitive contribution to our experience of it. Whether we can or cannot represent the external world as it is in itself is not a question that excludes that it is what it is solely in virtue of itself. But if we are going to know reality as it really is, our concepts have to reflect its

[4] Nagel (1974).

very nature. As Theodore Sider expresses it with approval: "The world has a distinguished structure, a privileged description. For a representation to be fully successful truth is not enough; the representation must also use the right concepts, so that its conceptual structure matches reality's structure. There is an objectively correct way to 'write the book of the world.'"[5] That is the God's eye view. Reality is not just what it is as a result of our true apprehension; knowing reality requires that human comprehension transcends its own experiential conditions by having a truth content that corresponds to this reality.

However, a metaphysical realist could in principle hold that the external world may be both structured and ontologically determinate, or unstructured and ontologically indeterminate, but whether it is the one or the other, it is what it is beyond any finite being's epistemic capacity of representing it. But a metaphysical realist claiming that the world is species-independent is not required to believe more than that. She is not forced to believe anything specific about the structure of reality. She may deny completely that the ultimate reality is what it is considered to be on the basis of our ordinary experience. Instead she may argue that the objective species-independent world is as the most advanced sciences tell us it is. The later formulation may be called the *metaphysical version* of scientific realism. This is the view Kant scornfully called transcendental realism. Whether the realist adopts a common-sense or a metaphysical perspective on scientific theories, he holds a view in which scientific theories tell us about a world hidden from our immediate senses: the world is furnished with different kinds of particles and forces impossible to see with our naked eye, or even conceived in our imaginations, and which do not possess the same properties as those that common sense realism ascribes to perceptual things.

Setting the various versions of *scientific* realism aside for a moment, what arguments can be levelled in support of metaphysical realism *in general*? Many will probably agree with Thomas Nagel when he points out that if we look at our history, we see that at some time our predecessors did not know, or were not able to conceive, aspects of reality that we know or can understand today.[6] Similarly, we can infer inductively from

[5] Sider (2011), p. vii.
[6] Nagel (1986), Chap. 6.

this fact that there are things we cannot now grasp, but our descendants will be able to in the future. From these observations, most people will accept an inference to the conclusion that there may be things we cannot comprehend at any particular time in the future and therefore never ever come to understand. Of course, the decisive factor is whether this implies that there are things of which we have no conception because of the sensory capacity we have and the way these things are, and not because we are not yet technologically developed enough to detect these things. Here the waters divide between metaphysical realists and common-sense realists.

On the one hand, the metaphysical realist would argue that some people lack a capacity to conceive of colors or sounds if they are born blind or deaf. And some people do not have the mental ability to understand advanced mathematical theories. Analogously, we can imagine that there are aspects of reality that nobody, in principle, will ever be able to experience in form of perception or observation. The common-sense realist, on the other hand, would dispute this argument by arguing that our thought cannot reach beyond the conditions for the possibility of human action and experience. If she is also a Darwinian, she may add that these conditions are themselves a function of adaptation by natural selection. We can make sense of the examples of the disability of the blind, the deaf, and the person with a low mental ability to see, hear or understand aspects of the world, only because we realize that other people may not lack the ability to know or conceive them. In other words, the Darwinian believes that the examples make sense since we who belong to *Homo sapiens* have a cognitive capacity by which these features are fully specifiable. The Darwinians maintain that we cannot claim to have a general concept of reality based on what we know or comprehend already and then meaningfully apply it to something that is totally incomprehensible.

The disagreement between the metaphysical realist and evolutionary naturalist boils down to a dispute about the objective nature of reality, not about the existence of an external world. This dispute cannot be addressed further until we know more about the boundaries of our thought and how truth relates to sentences expressing our thoughts. But Nagel mentions that in our notion of a universal or an existential quantification, the value of a variable does not need to be the referent of a specific name or description in our language.[7] The reason is that we already have a general concept of everything, which comprises both what we can name or describe, and what we cannot. Consequently, we can speak of "All the things we can't describe, imagine, or conceive of owing to our very nature."[8] For this claim to become a way out for the metaphysical realist, it seems as though he must admit that such a sentence can be true only if reality-in-itself consists of a negative fact that makes the sentence true. So long as the realist talks negatively about something that is known, say, "The Eiffel Tower is not made by wood," a statement like this does not require to be true a negative fact that the Eiffel Tower is not being made of wood. What makes it true is the positive fact that it is made completely of steel. If it is completely made of steel, therefore it cannot also be made of wood. In the case of a sentence concerning something we cannot describe, or imagine or conceive, the realist cannot state any positive fact that makes the negative sentence true. Thus, if this consideration is correct, it raises serious doubts about Nagel's claim that the general concept of reality, he applies to what humans cannot understand, is the same as the one he uses for what is conceivable by us.

The above account of realism as regards the external world has so far provided us with three possible realist positions: (1) Physical things that we experience immediately through our senses or experimental operations exist in some way or another irrespective of no human perception of them; (2) physical entities, including the laws of nature, which are not objects of direct sense experience, but which are postulated by our best scientific theories, are also real in the same sense as the objects of direct sensation, and not merely mental constructions; (3) the best scientific theories aspire to tell us what the ultimate reality is. We may associate these three positions with *common-sense realism*, *scientific realism*, and *metaphysical realism* concerning scientific theories. Common to both scientific realists and metaphysical realists is the fundamental conviction that their theories *represent* more or less the world as it is and therefore

---

[7] Nagel (1986), p. 98.

[8] Indeed, there are things I can conceive of, but not imagine, like multidimensional spaces, and there are things I can imagine, but not describe, like the taste of a pineapple. And there are things I can describe or conceive of but only inadequately imagine, like a chiliagon (Descartes' 1000-sided figure) or a bat's 'seeing' the world by sonar.

are true or approximately true. In contrast, the common-sense realist distances himself from the representational view of knowledge.

Nonetheless, it is not uncommon to hear an objection against any attempt to account for realism. The complaint is that realism specified in terms of commitments to a mind-independent world is obscured by metaphorical language. Is it possible to specify the realist's position further? Maybe not. But we draw the distinction from our personal experience between concocting a story by imagination and telling a story from memory of real experiences. Nobody, I assume, will deny that we are responsible for some of our beliefs and not accountable for others. There is nothing more to the claim that the external world is mind-independent, i.e., that it is *external*.

A possible supplement could be to say that realism with respect to the external world also involves a semantic commitment: If the world is not co-extensive with our cognitive resources, but is represented by our thoughts, then sentences about physical laws and objects are not reducible to sentences about mental states. The former type of expression has a meaning that cannot be translated into expressions of the latter type. This is the realist's *untranslatable thesis*. For example, according to common-sense realism, sentences about the external world are not translatable into sentences about sense data, and the truth of physical-object statements cannot be expressed in terms of the truth of statements concerning mental states or subjective experiences. I am not claiming that this semantic formulation is logically equivalent to the ontological formulation of the mind-independent thesis. What I am saying is that for the realist, the mind-independent thesis has to be associated with the untranslatable thesis to be intelligible. This holds for metaphysical realism as well as for scientific realism and common-sense realism. Even though the metaphysical realist admits the possibility of completely *unobservable* entities, he cannot claim without serious difficulties that this part of reality may be completely unknowable and still hold that our language can be used to describe such an invisible world. Though logically possible, it is difficult for the metaphysical realist to argue positively for the existence of a reality *an sich* and at the same time hold that the nature of this reality as such could be cognitively inaccessible in principle, because if the world-in-itself can be described by a language, it must be possible to check their truth value, and if it cannot be so described, it would be impossible to get to know it. The common-sense realist argues, however, that we are prevented from having knowledge of the nature of the things-in-themselves and that this implies that beliefs about ultimate reality cannot be expressed in physical-object sentences. Consequently, the realist concerning scientific knowledge can be characterized semantically as one who argues that: (1) The truth conditions of statements about theoretical entities cannot be reduced to statements about what we can perceive, but also if she is a metaphysical realist; (2) the truth conditions of sentences concerning laws of nature cannot be reduced to sentences about the physical manifestations of these alleged laws.

Based on the above discussion, we may characterize metaphysical realism as a general philosophical doctrine consisting of three components. First, there is the *ontological* component of the view: Whatever there is is what it is independently of our cognitive capacity to think of it. A real entity, whatever that may be, has an objective specificity and determinateness, or it lacks both, independently of our cognitive powers. The metaphysical realist is not forced to argue that determinateness holds good for reality as a whole. For instance, instead of maintaining that the future (and the past) is ontologically determinate, she could claim that the future (and the past) is ontologically indeterminate or simply unreal. Likewise, she could argue that some substances are vague or fuzzy entities that have indeterminate attributes. This leaves, apparently, the metaphysical realist with two different options concerning the nature of the species-independent world. First, she could hold that everything real is ontologically determinate in the sense that it has concrete specific attributes; second, she could hold that at least a part of what is real is ontologically indeterminate in the sense that it lacks determinate actuality and attribute specificity.

Second, there is the *semantic* component of metaphysical realism: the meaning of statements about reality must be analyzed by reference to the notion of truth conditions, the specification of which in principle may reach beyond any possible empirical justification. A sentence is true or false independently of whether or not we have any means to verify or ascertain its truth value. What determines these truth conditions is an alleged objective relation between a statement and the world that makes it true. A set of descriptions or a scientific theory is true if, and only if, it is related to the world in a way that corresponds to the world as it really is. Although reality in itself, according to the metaphysical realist, exists as it does entirely detached from our cognitive capacities, it is generally assumed that those subject-predicate-object statements of ordinary language and/or scientific statements we use in our communicative discourse refer to such a species-independent world. Thus, claiming that a species-independent world exists is associated with a thesis that true scientific accounts concern an objective reality, regardless of our sensations, opinions, or emotions. An important consequence of this thesis is that statements about the world are not reducible to statements about anything else, especially not to statements about our sensory experience or mental states of the subject.

The third element is the *epistemic* component: we may have objective a priori knowledge of the world as it is in the sense that we make a rational analysis of what counts as *the* correct theory of reality as a whole. Knowledge in this objective sense is beliefs, which all rational beings would agree on if they had the capacity to justify their certainty. Thus, the metaphysical realist maintains that objective knowledge exists in the form of invariably true propositions and scientific theories. In other words, propositions and theories concerning the reality-in-itself are held to be true independently of whether we have experientially proven, or might prove, them or not. As Karl Popper states this position: "Objective knowledge is knowledge without a knowing subject."[9] This is knowledge as God would have it.

Since reality *an sich* for God would be one with his understanding of it, God does not, according to such a viewpoint, need reliable methods to prove God's possession of objective knowledge. The world-in-itself would be inseparable from God's knowledge of it, or reality *an sich* would at least be congruent with his conception of it,[10] for God as an infinite mind would not be bound by a distinction between the subject and the object. But finite human beings, in contrast, need reliable methods to determine whether or not their mental representations such as scientific theories are in accordance with reality *an sich*. Thus, the fourth element of metaphysical

[9] Popper (1972), p. 109.

[10] Only the latter would be the view of traditional Christian theology. The pantheistic or panentheistic view that reality exists in God is a heresy for traditional Christians, since they also hold that there was evil in the world.

realism is the *methodological* component: in the right circumstances ordinary people or scientists are able to provide warranted judgments about the truth of all kinds of beliefs regardless of whether they are about observable or unobservable entities or are formulated in terms of singular or universal sentences. This is due to the fact that some rational (perhaps not empirical) methods or procedures exist such that their application yields a true belief of the world as it is. Beliefs about the reality as a whole, according to the metaphysical realist, are ascertainable by rational means: nevertheless, there are other procedures that, when followed, yield only good, and not certain, grounds to believe that something is objectively the case. Such a procedure provides us with a rational method by showing that the appropriate statement is likely to be true or false.

In order to defend this position, the metaphysical realist is bound to explain what kinds of facts make statements about the world-in-itself true. He must give us a metaphysical account of how the truth value of statements about ordinary entities, about unobservable entities, and about abstract entities are ascertained. Furthermore, the metaphysical realist must explain how we can have different epistemic access to ordinary things, the realm of an unobservable reality, and universal, modal and necessary truths. He must point out which truth-conducive procedures of inquiry are at our disposal for gaining such knowledge. He must also identify under which circumstances we can know that the truth conditions of the one and only correct theory are in fact fulfilled and, in general, what conditions have to be fulfilled for a meaningful use of the sentence in question. Indeed, the metaphysical realist's position becomes precarious if his analysis of the truth conditions means that metaphysical facts lie beyond the empirical domain of direct or indirect evidence. Indeed, the metaphysical realist could deny that he is committed to all three components. Instead, he could argue that a world-in-itself exists, but apart from this we cannot say anything about it. In this case he would not say more than what Kant did.

4.2 Conceptual Frameworks and External Commitments

Having laid out these various forms of realism, it must be emphasized that some philosophers see themselves as both scientific realists and empiricists. This is true of Karl Popper and Hans Reichenbach to mention only a couple. Others, like Bas van Fraassen, call themselves empiricists and scientific antirealists. Whether one prefers to call oneself a realist or an antirealist is more or less inconsequential, so long as one holds most of the common-sense realist's presuppositions as one's own. More important than such labels is that a given view is characterized unequivocally and exhaustively. However, there seems to be a tendency among those empiricists who consider themselves as epistemic optimists to accept that there are some methods that can provide us with a rationally warranted belief in the claims of science, methods, that is, that make scientific statements regarding things beyond the directly empirically accessible sufficiently probable that it is rational to accept them. In contrast, epistemic pessimists argue that there are no reliable procedures of inquiry yielding a rational warrant for believing the truth of scientific theories.

The opposition to realism with regard to theoretical entities beyond the empirical world has traditionally been championed by the instrumentalist doctrine that theoretical concepts are merely heuristic tools for organizing the scientist's observations. Instrumentalists take a nominalist stand on theoretical entities. Common names and natural kind terms of invisible entities do not refer to anything in reality; hence, statements about these entities should not be considered literally true. All concepts of "unobservable" things, events, and properties are nothing but logical constructions from "observables." Accordingly, the central thesis of this view is that imperceptible things like forces, fields, atoms, molecules, genes, and viruses are not real and that the names of these things proclaimed are merely a unifying designation of concrete experimental results. This contention leads to the claim that scientific theories containing sentences about such imperceptible things do not express proper knowledge; instead, they are inference schemes that can be utilized for predictions of future experiences on the basis of past experiences.

Instrumentalism is an ontological position about theoretical entities closely associated with the application of phenomenalist constraints on what can possibly exist. Only things with which we are directly acquainted can be said to exist with any justification for phenomenalism. Embracing such strong epistemic requirements on ontology, instrumentalism can be regarded as a form of non-cognitivism about what we cannot directly perceive. Similar non-cognitivist views have been asserted within other areas of human cognition: Discussions about the reality of tenses, moral values, causality, probability, and possible worlds can in many cases be seen as a continual battle between realists and nominalists. Therefore, the question we need to ask is whether the instrumentalist has better arguments against the existence of theoretical entities than those of the phenomenalists against the existence of ordinary physical objects.

The language of science is full of terms that refer to invisible entities and properties. Therefore, one seems to be ontologically committed to entities and properties that we cannot perceive, unless the instrumentalist can prove, for instance, that all sentences concerning them can be translated without loss of meaning into sentences of a language in which each and every term concerns perceivable objects. Few instrumentalists, other than operationalists, would argue that a given theoretical sentence has the same intension as any set of observation sentences, that the truth conditions for a sentence "*X* is *F*" containing terms for a so-called "unobservable" object *X*, and a similarly "unobservable" property *F*, are identical with the truth conditions of an appropriate observation sentence, or a set of sentences, "*Y* is O_1, O_2, O_3, O_4,…, O_n," which contains only the terms for an "observable" object *Y* and the "observable" properties *Os*. An instrumentalist does not have to argue that these two sentences necessarily have the same meaning.

As an alternative the instrumentalist could say that he does not claim the synonymity of such sentences but merely considers them to be coextensive. One way to vindicate such a consideration is to do as Frank Ramsey proposed and substitute existentially bounded variables for predicates and names. He proved that all theoretical predicates of a theory, i.e., terms referring to 'unobservables,' can be treated as existentially quantified variables such that the axioms of the theory link the predicate variables to each other, and a dictionary links them to the observables.[11] The result of this so-called "Ramsification" is that all problematic predicates are eliminated but the structure and observational consequences remain unaltered. If the so-called "Ramsey sentence" is true, it tells us that to which we are ontologically committed. Therefore, Ramsey sentences have been used in the attempting to get rid of theoretical terms and replace them with observational terms. In fact, this was not Ramsey's own purpose. Rather, he used his method to define the observational terms of observational language in terms of the theoretical terms of theory.

The instrumentalist disapproval of the fact that in order to be true the language of science presupposes the existence of things we cannot observe by the naked eye is only one of two challenges to the ontology of "unobservables" with which the realist must deal. It is simply not enough for the realist to prove that the language of observables cannot express all our scientific beliefs. Instead, the common-sense realist may rejoice that it is part of our biological heritage to argue that things we can use, manipulate, and handle are real things that do not go in and out of existence with human thought of them. In this respect, invisible things are no different from visible ones. The other challenge arises from the fact that to be true the language of logic and mathematics, for instance, apparently requires the existence of abstract entities. In Peano's arithmetic we seem to be committed to holding that natural numbers exist as real entities apart from figuring in mathematical discourse, and in the Zermelo-Fraenkel set theory we have the same obligation toward sets.

As Rudolf Carnap once pointed out, whenever we adopt such linguistic frameworks we are ontologically committed to believe in the reality of numbers, sets, propositions, and so on.[12] But he also made the restriction that ontological existence claims outside a particular linguistic framework are devoid of cognitive content. By this he meant that we have no cognitive means at our disposal to determine whether or not the truth conditions of such sentences are satisfied by some non-linguistic entities. He argued that whenever we wish to talk about some kind of being, we must do so within a linguistic framework. Such a framework is constituted in virtue of (1) a

---

[11] Ramsey ([1930]1990).

[12] Carnap ([1950]1958).

set of concept definitions, (2) some principles for governing the syntax between these concepts, and (3) some principles for testing the truth values of statements within the framework. In the case of a rational (as opposed to an empirical) framework, (2) and (3) are coextensive. The commitment is '*internal*' with respect to the linguistic framework employed by any given theory. Thus, a sentence like "Is there a prime number greater than 5?" can be answered within the theory of numbers by saying that at least one number exists of which we can say that it has the property of being prime and greater than 5. Carnap also argued that no metaphysical question about the existence of numbers as a whole can be answered "outside" this framework; that is, no existence claim has a truth value with respect to the world independently of some linguistic framework. The request for the truth value of existential claims independently, or "external" in Carnap's terminology, to a linguistic framework is cognitively meaningless. When we ask if something *really* is, we are asking a question that goes beyond the conventional criteria for establishing whether something is. It is an '*external*' question to which there can be given no real meaning because it concerns reality considered outside a linguistic framework.

What Carnap did was to maintain that mathematical theorems are meaningful when we have specified the syntactic and semantic rules for talking about numbers, including what it means to say that mathematical statements are true, but without committing ourselves to saying that mathematical objects exist independently of the language in which they are specified. Any attempt to legitimize the reality of any entity posited by a linguistic framework in other terms than empirical evidence is self-defeating because their existence is the product of the linguistic framework. However, Carnap may be criticized for setting up his distinction based upon assumptions that he regarded as philosophical bedrock, but which no philosopher shares today. Those assumptions were some form of verificationist principle of meaning connecting meaning and experience, an extensive form of linguistic conventionalism overstating our liberty of choosing a language and the classical analytic-synthetic distinction.

Anyhow, Thomas Hofweber has convincingly defended the distinction by arguing that a standard question like "Are there numbers?" can be given two readings of the quantifier contained in the statement and therefore be provided with two sets of truth conditions.[13] One he calls the inference role reading, another he names the domain conditions reading. The first is the trivial reading allowing a competent speaker to infer from one sentence to another and therefore knowing how to answer the question, whereas the second is not at all trivial. What we ask about here is whether there exist numbers, or what have you, external to and independent of the language of our discourse. Hence, the external question just as much as the internal question is a question about facts.

In fact, the plausibility of Quine's famous dictum "to be is to be the value of a bound variable" hinges on a similar dichotomy between internal and external commitments.[14] Indeed, this claim may seem very odd, since Quine personally rejected the significance of the division between internal and external existence.[15] Existence is what existential quantification expresses. Thus, the ontological commitment of a given theory can be found by identifying the entities over which the quantification of the theory is made. Internal questions such as "Is there a prime number greater than 5?" are subclass questions, whereas external questions such as "Are there numbers?" are categorical questions. However, Quine thought that categorical questions can be reformulated as subclass questions. For instance, we could interpret 'Are there numbers?' as a subclass question: "Of all the objects are any of them numbers?" So the distinction is only relative to a particular framework. Quine himself might have approved Carnap's distinction if it had been possible to make a clear-cut distinction between analytic and synthetic sentences, but it is well known that he did not think this was possible. However, Quine's conclusion can be put into question: logic alone does not have the resources to answer any questions concerning the existential status of the objects over which the variables range. In fact, as Hofweber also points out, non-referential words, for instance, non-referential number words, are congenial with quantification over numbers used in their internal reading. The objects and the variables are logically distinct, but are the objects real or merely parts of linguistic constructions for the purpose of interpreting the variables?

[13] Hofweber (2007).

[14] Quine (1969), pp. 91 ff.

[15] Quine ([1951]1966).

To settle the issue we must look for answers "outside" of the linguistic framework.

Putnam's internal realism rides on the same ticket as Carnap: "'Objects' do not exist independently of conceptual schemes. We cut up the world into objects when we introduce one or another scheme of description. Since the objects and the signs are alike internal to the scheme of description, it is possible to say what matches what."[16] Similarly, van Fraassen's claim that we are committed to accepting our best scientific theories, though we cannot establish their truth, rests on the distinction between internal questions about what the theory postulates as an interpretation of its formalism and external questions about whether or not the posited entities are real. A common-sense realist of the Darwinian sort would say that such external commitments can be explained in virtue of our innate capacity of distinguishing language from what it deals with. None of our hominid ancestors would have survived if they had not been capable of separating representation from what is represented. Were our distant forefathers and foremothers, as a result of selection, tuned in only on a warning signal and not also on the danger it could be associated with, we would not have been here.

The metaphysical realist's (or, for that matter, the metaphysical anti-realist's) commitments are in conflict with Carnap's thesis that external questions about the ultimate reality have no cognitive satisfactory answer: The reality of abstracts such as numbers, sets, possible worlds, and propositions is a question about what *really* exists (or what *really* does not exist) independently of any conceptual framework. Solving this disagreement is pressing, as the following analogy shows. Mathematicians have constructed different geometries based on distinct interpretations of the famous parallel postulate. These different geometries as conceptual frameworks are all internally coherent and consistent like all good theories of some metaphysical entity. But which one of these alternative geometries is the one that truly represents that physical space is an external question? In science it becomes an empirical issue which one of these geometries gives us the best representation. However, within metaphysics the best empirical representation is not necessarily the true representation,

[16] Putnam (1981), p. 52.

but any appeal to coherence, consistency, intuition, or inter-theoretical connections does not help us to answer the external question.

A similar external commitment holds for the scientific realist who argues that the aim of scientific theories is to give us not merely the best but a true description of the empirical world. Thus, he must be prepared to argue for the correctness of the assumption that space, time, atoms, quarks, fields, and so on exist objectively regardless of our way of conceptualizing this world. The scientific realist is forced to show that her beliefs about the empirical world can be warranted in some ways other than just by appealing to a given linguistic framework.

For instance, classical mechanics relies on everyday concepts like solidity, motion, and position in the observational description of macroscopic objects. But the usually crude determination of these attributes was not entirely satisfactory with the recognition of the Renaissance that they could be measured and therefore become represented by the objects of mathematics. They could be turned into quantities. From that time on, a precise determination of their magnitude would involve instruments. Rulers, clocks, and levers were the basic instruments, and thereby mechanics got a new set of observables that were instrument readings that could be expressed numerically. Such pointer readings must be connected with mass, position, and velocity through operational rules: meter sticks gauge the scale of distances, clocks record how much time has elapsed, levers can be used to measure the weight of masses in a gravitational field, and velocity can be defined as uniform distances covered by equal times.

Newton's mechanics ascribes imperceptible properties to perceptible entities. The ascription can be done through those of their properties we can experience. Quantum mechanics, however, deals with theoretical objects that cannot be objects of direct perception; hence, none of their properties can be attributed to them on the basis of our visual acquaintance with any of their other properties. Nevertheless, both William Craig[17] and Carl Hempel[18] have independently shown that with respect to any axiomatizable theory, it is always logically possible to construct an equivalent theory that entirely leaves out theoretical terms and expressions and

[17] Craig (1956).

[18] Hempel ([1958]1965), pp. 173–226.

replaces them with observational terms and expressions. Thus, theoretical terms are construed as meaningless auxiliary marks that serve as inferential devices between observational statements. Indeed, choosing a theory without theoretical terms has severe costs, such as loss of explanatory power, simplicity, and heuristic fertility. A common-sense realist is not obliged to prefer such a language as the one that can express what we can know. She may argue that human thoughts have been set up such that things leave traces, and whenever we have experienced enough of them we have to believe that something has left them. In particular, such beliefs are inductively justified if in other situations we have been able to interact with what causes these traces.

The realist position concerning invisible entities is very often identified with the thesis that theories that are considered to be the best at present are closer to the truth than earlier ones and that the central terms of our best current theories are genuinely referential. This identification means, of course, that the truth of theoretical sentences about invisible objects and attributes is not reducible to the truth of a finite set of sentences about empirically accessible things and properties, or, in other words, a Ramsay sentence is not synonymous with the original theoretical sentence. As a reason for her position, the realist may point out that only if modern scientific theories are regarded as approximately true can we explain their predictive success. However, a realist may seek a stronger commitment than to those objects which science may posit. Science manages, she may argue, to describe laws and structures far beyond our ordinary experience. The physical world consists of both pre-given entities as well as pre-descriptive laws of nature. And, according to her, the aim of scientific theories is to give a literal and objective description of such a world, and its present success can be seen as a token of the performance of these efforts. The realist of this metaphysical sort holds that science eventually secures increasing knowledge about the world as it is in itself and hence knowledge about the basic structure of reality. However, being an evolutionary naturalist myself, I could imagine holding that the entities, their properties, and relations are real and not holding that laws have some sort of objective reality beyond the things whose behavior they describe. It seems that you could be a realist with respect to entities but a nominalist with respect to laws.

4.3 Theory Realism

In the preceding chapter I argued that an evolutionary naturalist can account for our beliefs concerning the existence of objects that we cannot see with the naked eye but that can be observed by instruments. So the question is to what extent the scientific realist needs to assume metaphysical realism concerning things we cannot even observe by instruments before his assumptions clash with the evolutionary naturalist's common sense. The evolutionary naturalist would exclude these genuinely invisible entities from our ontological commitments. The problem we have to address in the remaining part of this chapter is therefore how much of this realist's interpretation of scientific theories can be defended based on evolutionary epistemology.

Usually, the scientific realist is committed to a belief in a world of invisible entities governed by laws of nature, which it is the goals of science to discover. But what counts as theoretical 'entities,' 'natural laws,' and their 'properties'? How many or how few of the scientific terms employed in theoretical explanations stand for real objects? Apparently, it varies from one science to another which entities or quantities we regard as observables as well as from one explanatory context to another in the same discipline. Here it would be appropriate to remark that one may "pick and choose" with respect to which theory one interprets 'realistically.' Even in classical macrophysics there are clearly many non-referential terms like "frictionless pulley" or "perfect vacuum," or "black body," which we usually call "idealizations." One may be a realist about terms employed in most of the explanations of quantum physics but less enthusiastic about realistically interpreting many of the "unobservables" of contemporary cosmology. Furthermore, one may be totally skeptical with respect to a realistic interpretation of the explanatory vocabulary of Freudian psychoanalysis.

In general, macroscopic objects and events can be seen by the naked eye, and their visual properties like size, shape, form, solidity, colors, position, and motion are what distinguish them from each other. Human beings are aware of these attributes as a result of selection and adaptation. However, in the Renaissance the physicists began to attribute a new kind of properties to physical objects based on their wish to employ measurable quantities. These enabled quantitative prediction that could be more easily checked with observations, thereby justifying the theories employed in the explanations. Although such quantitative attributes were of a human construction, they were considered to be nature's primary and real properties. Some of the visual properties had to give away in what was intended to be an 'objective' description of physical objects. These non-quantifiable attributes were treated as secondary and mind-dependent properties because they could not be replaced by measurable quantitative properties.[19]

In physics, experience tells us that qualitative attributes do not play a role in the description of the kinematical or dynamical behavior of physical systems. In classical mechanics, for instance, an object's position, velocity, rotation, and acceleration are the intended properties that are immediately accessible to the senses. Its mass is also a property we sometimes experience directly as the solidity of matter and feel by the gravity. But transformed into a quantity on a gravitational force field, it can be measured by the weight. Basically all measurements are position measurements since positions and changes in position are the only quantitative properties we are adapted to see. All other mechanical entities and properties like force, momentum, and kinetic energy are defined as not directly observables; however, they can all be specified in terms of directly observables: $F = ma$, $p = mv$, and $E = \frac{1}{2}mv^2$. Classical mechanics ascribes such "non-observable" properties to all physical objects on the basis of "observable" ones. But the scientific realist would say that these "unobservable" properties are something over and above the various relationships between the "observable" properties. In the last chapter we recognized the arbitrariness of this old distinction. Properties such as force, momentum, and energy may be imperceptible but not unobservable. Even though these attributes are not directly represented by a corresponding pointer reading observable, we do have observable access to them whenever the pointer reading observables provide us with a second-order belief of the value of these properties. It is part of the meaning of these terms that the changes in the position of a system having

[19] Well, being so 'treated' was hardly universally agreed upon. The very arbitrariness of this distinction, relying in turn on the arbitrariness of 'measurable' led to its denial (Berkeley) and the onset of the idealist worldviews that characterized the Romantic period. We cannot really ignore the whole Kant through Hegel excursion, folly though it may be.

a certain mass constitute conclusive evidence (within a theory) that the system has a definite momentum, a specific kinetic energy, or is acted upon by a specific force.

In his defense of realism, Michael Devitt presents us with the following way to characterize this ontological position: A person *P* is ontologically committed to believing in the existence of an object *a* (or a property *F*) in uttering assertively a sentence token *S* if *a* (or *F*) must exist to make *S* true. Though Devitt will not deny the validity of this semantic criterion (and we have just seen in the discussion of Carnap's linguistic frameworks how such a criterion fail to meet external commitments), he believes that there is another, more basic criterion, according to which a person is so committed if, in asserting *S*, that person says that *a*, or an *F*, exists.[20] The first criterion requires that we possess a semantic theory for *S* to tell us what must exist to make *S* true, before we can say anything about a person's commitments, whereas the second criterion merely presupposes that as speakers of a certain language we understand *S* if, and only if, we know what ontological commitments a person has. If someone asserts "The electron is an atomic particle," this sentence is not true unless there exists something to which "electron" refers and to which "atomic particle" applies. But Devitt claims that the commitment of this statement to the existence of electrons and atomic particles is the same as the one following from the assertions "The electron exists" and "Atomic particles exist." Period!

I think Devitt's argument is correct only if it is taken to establish that no semantic theory is needed to know what "existence" really means. The word "exists" in a sentence like "The electron exists" does not have a meaning distinct from its meaning when we claim that the electron must 'exist' for the sentence to be true. Had there been any difference between its meaning in the object language and the meta-language, we could decide to replace the meaning in the object language with its meaning in the meta-language, or, if not, we might be involved in an infinite regress. But the fact that there is no difference leaves us without an argument for the conclusion that our commitments are external to the linguistic framework. Moreover, if the Craig-Hempel thesis holds, then since any theoretical sentence can be proved to be coextensive with a set of observation sentences,

[20] See. Devitt ([1984]1991), sec. 4.6.

the realist is deprived of a strong reason to claim that our ontological commitments are external to the theory. If a theoretical sentence expresses a fact that can be translated into a certain appropriate set of simple observation statements, we would be able to do away with any reference to "unobservables." Thus, why are we being justified in believing that imperceptible entities and properties are real? Does the theory realist have something better to offer?

The theory realist sees the success of science as strong grounds for her thesis that the theories of developed sciences are typically approximately true. This success is also taken as evidence for the contention that theoretical terms within our best theories refer to whatever they are supposed to refer to. Sometimes it is even said that realism is the only conceivable explanation of why science has been so successful, because the prediction of observable phenomena would be a cosmic coincidence or a miracle if theoretical terms did not refer to real entities.[21] If realism were not true, it would be especially incomprehensible how new and unforeseen phenomena can be predicted by a theory that it was not created to predict. The discovery of the element hafnium succeeded its prediction on the basis of Bohr's reorganization of the periodic system according to physical features of the atoms. As a consequence of his relativistic theory of the electron, Dirac predicted the existence of a positive electron before Anderson discovered it. Pauli suggested the existence of an escort particle, the neutrino, as an explanation of the continuous spectrum from beta decays; its existence was not directly confirmed until many years later. The exchange of virtual mesons in a nuclear field was an essential part of Hideki Yukawa's theory of the strong nuclear force before these particles were discovered about ten years later. The W bosons and the neutral Z meson were first detected after they had figured for a while in Steven Weinberg and Abdus Salam's theory amalgamating the weak and the electromagnetic force. All such examples make it highly unlikely, the realist contends, that theoretical terms making these predictions possible do not refer to real entities causing those phenomena that are perceived.

The theory realist, however, also adduces other arguments for her thesis that theoretical terms may refer to something objectively real and that

[21] See Smart (1963), p. 39, and Putnam (1978), pp. 18–19.

therefore we are ontologically committed to "unobservable" entities in an external rather than an internal sense. In searching for a systematization of their data with the purpose of explanation and prediction, scientists need (because it is inherent in the nature of scientific explanation) to operate with hypothetical entities that are not directly observable. The hypothetical method, which was at first scorned (particularly atomism), came to be accepted because theories employing such postulations were more explanatorily and empirically successful than rival theories that stayed within a purely empirical vocabulary. As long as scientists confine their explanatory resources to perceptible entities, the realist argues, they are able to formulate merely empirical generalizations, but generally scientists are not content with the amount of integration that empirical generalizations alone deliver. What they want is a further integration of laws that bases itself on a small number of scientific principles, something that requires a further unification and development of concepts covering a broader domain of experience. Therefore, the way to pass beyond empirical generalizations must be accomplished by introducing more general concepts not corresponding to anything directly observable. And, says the realist, scientists eventually get a better and better grasp of the world through their acquaintance with these principles as they become able to expose the laws or mechanisms underlying the phenomena.

But how can this be an argument for the reality of natural laws or invisible entities? The realist argues that the ontological commitments entailed by our most successful scientific theories are justified because they make possible an integration of concepts. Scientists seek such unification only partly because of pragmatic reasons; that is, they want to work with as few conceptual tools as possible. A more important reason is that most scientists believe that their concepts reflect something in the world. So if they can manage to reduce the number of the general concepts in their description of a certain domain, they have reason to believe that at least that part of nature has been described in its most basic form, which the realist may then take to be as it is in itself. This realist line of thought is that whenever science is capable of describing the world with all its difference and complexity, given very few concepts, it is most likely to be true of the world as it ultimately is because these concepts have dissolved the complexity into its most simple constituents.

This argument, however, suffers from two serious shortcomings. The first one is due to the fact that the conclusion is not consistent with the history of science. Many abandoned theories, once used to explain an entire domain of experience in virtue of few general concepts, are no longer taken seriously, and the entities they postulated form no part of any contemporary ontology. Think, for instance, of the Aristotelian theory of motion. During the period it was widely accepted. It seemed to give a coherent account of our everyday experience of motion based on a few simple concepts. Vertical movement was considered dependent on the gravity of the body; dense things like rocks and water went downwards, more ephemeral things like air, vapor, and fire, upwards. Horizontal movement of a wagon, a stone, or an arrow required the presence of a moving force in the form of oxen, horses, or manpower. All other motion could be described as a combination of these two fundamental forms. Likewise, the ancient idea of the world as built up of the four basic elements, earth, water, air, and fire, contains many fewer elements than any contemporary theory. Therefore, it is at least doubtful that today we have reached the right categories corresponding to reality once and for all just because we have been able to isolate a few concepts for explanatory purposes. The argument shows only that we always are *internally* committed to those entities and properties that our currently best theories presume—it cannot prove that we are *externally* committed to such things.

The second objection is even more fatal to the realist's argument. What reason could justify asserting that a scientific theory with fewer concepts is more likely to be *true of the world as it is in itself* than one with more concepts? The idea of an underlying unity beneath the diversity of phenomena is as old as philosophy itself, having been a primary motive of the Milesians' quest for an *arche*. The notion was strengthened in medieval times by the Christian doctrine that a perfect Deity would do nothing in vain, so He created the cosmos with as few as possible basic entities. This notion was further entrenched by the axiomatic development of mathematics and logic, which again began with the Greeks (Euclid) and has persisted right through the logistic program of Russell and the logical atomists. So I would tend to think that the fact that science is animated by such goals as unification and simplicity has a historical origin different from serving to 'justify' (in the eyes of realists) postulating hypothetical entities. Today there are really no metaphysical grounds for believing that the world should consist of only a few basic entities instead of multiple such. Similarly, nothing implies that these entities must have fewer properties rather than more. Even if we grant the scientific realist the existence of such assumptions, it is impossible to see how that could help him to establish his belief that there are those entities or properties that a certain scientific theory prescribes, for such a theory may turn out to be too simplistic in its assumptions about the basic number of entities or properties constituting its domain. Historical evidence tells us that frequently theories start out by postulating very few entities and properties, but eventually have to adopt a lot of conceptual extensions in order to cope with increasing experimental evidence for entities or properties not included in the original version of the theory. Clearly, we do not particularly want a theory that posits superfluous entities or properties. But rejecting superfluousness is not the same as embracing simplicity.

From a Darwinian point of view, it is not surprising if we are programmed to be epistemically attracted to the simplest solution. This is when inductive learning works best, it makes our reasoning easier to grasp, and it saves time to work with the simple rather than the complex. But does simplicity tell us something fundamental about the world, beyond the fact that as a mental capacity it has been beneficial for us?

In addition to the reasons discussed above, further arguments have been advanced to support the realist claim that there are objective counterparts to postulated entities and properties. Closely related to the unification argument is the question of abduction or inference to the best explanation. Against the antirealists, realists argue that only some scientific theories operating with "unobservable" structures and mechanisms can explain all relevant facts in a coherent and convincing way. Thus we have grounds for concluding that those theories are able to do so tell us how the world really is, or at least approximately how it is in itself. However, we have to distinguish between at least two kinds of claims that might motivate embracing this inference to the best explanation. On the one hand, the realist may hold that an inference to the best explanation leads us to the objective laws of nature, in which a case he may be called a realist concerning scientific theories; on the other hand, he may embrace the claim that the inference shows what is the most likely entity causing the effect, and in that case he could be said to be a realist concerning scientific entities.

Induction, as well as inference to the best explanation, plays an important role in formulating appropriate theoretical laws of science. But, again, it is not strange for a Darwinian that such methodological capacities play an important role in scientific reasoning. We are born with a capacity of inductive learning. Adaptation has favored those who could form more accurate general assumptions based on relatively few encounters. This has, and has had, an advantage in the struggle for fitness. But the realist will have a hard time defending the view that inference to the best explanation is a reliable guide to objective truth. Historically, this inference has fallen far behind the production of infallible knowledge, and we have little basis for believing that the situation will change in the future. What is considered to be the best explanation at any given time is whatever theory or assumption that seems to cover all chosen known phenomena in the most satisfactory way.

The theory realist may attempt to be modest, saying that the inference to the best explanation only provides us with good reasons for an explanation to be likely more true. One may wonder, however, how to establish such likelihoods other than by saying that a theory is in empirical agreement with all phenomena considered being relevant at a given time. A correlation test, for instance, provides us with a measure of how good the correspondence is between the observed values and the expected values a given hypothesis predicts. Thus, if the measure of the likelihood is nothing but this external virtue, the realist must face the serious question of the empirical underdetermination of theories. Usually, though, the scientific realist will trade on internal virtues of a theory, like simplicity or fruitfulness, as the characteristics of the best explanation. But how can such internal virtues establish that the "unobservable" entities are real regardless of the conceptual framework?

The 'conceptual framework' is pertinent to the *description* a theory provides of the imperceptible entities or properties, but not to the *reality* of those entities. I am not arguing that it is inconsistent with realism to say that theories employing different conceptual frameworks may provide different descriptions of one and the same entity, without in any way impugning the reality of that entity, for example, phenomenological thermodynamics and statistical mechanics. I can well imagine a race of aliens that is intelligent enough to develop a scientific description of the world and that these aliens are referring to a similar world as our science. But I would be utterly amazed if they employed exactly the same conceptual schemes as we do in our science, especially if their sense organs were of a different sort from ours. Of course, conceptual schemes are tools and so relative to the scientists, their culture, and their goals and worldview. So what I attempt to argue is that the scientific theories may describe reality as it appears to the scientists but that it is impossible for the realist to argue that some of these different theories describe the world as it is in itself.

Simplicity and unification will not justify the realist's assumption, if he harbors such an assumption, that by and large theories come closer and closer to describing how the species-independent the world really is. But perhaps coherence and consilience might? It might be argued that the idea of a world-in-itself can be associated with the conception of everything being consistently connected with everything else. From this conception we might infer that a hypothesis capable of explaining the facts corresponds better to the world as it is in itself if it agrees with other hypotheses than if it does not agree with any. Newton was able to reconcile the laws in the heavens with those on earth. Kepler's laws and Galileo's could be derived from his laws of motion. However, a scientific realist may have a problem with such an argument since a true hypothesis may or may not cohere with most other assumptions considered to be true. Typically, a theory realist will subscribe to a correspondence theory of truth right as a matter of first principles. Assuming this theory, Aristarchus' heliocentric hypothesis was approximately true but inconsistent with Aristotle's physics and the majority of other astronomical hypotheses then accepted. When the view concerning the truth of these other hypotheses eventually changed, Aristarchus' hypothesis came into agreement with the majority of the newly accepted assumptions. Thereby physics got closer to reveal how the species-independent world really is. So the realist could argue that a claim is not metaphysically interesting, even if it is true, before we have independent, theoretical warrant for believing it. And he could continue by saying that so long as the hypothesis is not coherently connected with other commonly accepted assumptions about the world, it is not independently justified as true of the world as it is.

Also, the realist could emphasize that a hypothesis not only has to agree with other confirmed hypotheses to be closer to the truth than its alternatives. It also has to agree with those ontological principles forming the arrangement of the world, one of which I have previously named "the principle of the unities of time, space, and cause" after criteria of the classical drama.[22] For instance, the realist may argue that an explanation has an a priori probability of being true if it accounts for a certain phenomenon in terms of other phenomena with which it is spatially and temporally connected, all of which must fit into the same ontic scheme of categories that can possibly enter into a causal relation. Nobody, to put it vividly, would dream of explaining today's hole in the ozone layer over Antarctica by the assassination of crown prince Franz Ferdinand in Sarajevo 80 years ago, because we regard such an explanation as entirely irrelevant. And the reason for this claim of irrelevancy is that the explanation suggested does not respect the unities of time, space, and action. But supposing we could show that WWI led to a process of environmental degradation (and social change) that ultimately brought about the meteorological factors that caused the hole in the ozone layer, then citing the assassination in Sarajevo would not be irrelevant to explaining this phenomenon. Thus, the realist must supply arguments that justify assuming such an ontological principle and that therefore show that coherence with this principle is necessary for an objective description of reality. The fact is that what we can imagine *a priori* is due to our innate cognitive capacities and ruling out *a priori* certain phenomena as possible causes of other phenomena is a hazardous business beset with problems.

Another way of looking at the attempt to justify realism by inference to the best explanation is to say that such an inference leads us to those entities that are causally responsible for the observed phenomena to be explained. By assuming that real "unobservable" entities are causally responsible for what we can observe, realism yields the best explanation of why these physical phenomena occur in a lawful way. They do not just "pop up" by mere chance but are the observable effects caused by the behavior of "unobserved" entities. A theory that explains a wide variety of different phenomena according to a common cause is also bet-

[22] See, for instance, Faye (2002), p. 93.

ter than one that explains the same phenomena according to different independent causes. For example, as Wesley Salmon has pointed out, the determination of Avogadro's number, i.e., the number of molecules in a mole of any substance, was the decisive achievement in convincing the scientific community of the reality of atoms and molecules. What is crucial is not so much the fact that Jean Perrin succeeded in achieving a precise experimental value of Avogadro's number as the fact that, within a few years, he and others reached the *same* number based on several independent methods and carried out on a variety of different phenomena. Among those phenomena were Brownian movement, alpha decay, X-ray diffraction, black body radiation, and various electrochemical effects. Thus, ruling out the question of a 'miraculous' coincidence, this remarkable agreement among the results of experiments, which seem to be quite independent of one another, can be taken as strong evidence of the hypothesis that behind the different phenomena there is a real common cause of their appearances.[23]

Nevertheless, once again the history of science seems to teach us another and different lesson. As long as the discussion is kept on the empirical level, there are historical cases where theories were regarded as the most prolific explanations available, but where this explanatory success was not enough to establish the reality of the entities proposed. The theories of phlogiston and caloric are just two obvious examples. Apart from this fact, the antirealist is always in a position to argue, as Bas van Fraassen does, that a case of the type Salmon mentions merely shows that our best theories are empirically adequate.[24] Such a case does not by itself establish the general philosophical point that our theories of molecules have to be true or that molecules are real. This conclusion does not imply that modes of reasoning like consilience, coherence, and inference to a common cause have not shown their biological value in organizing sensory information. For instance, any organism that could recognize some particulars as all of a specific kind would be better off than those that did not have this capacity, simply because these organisms could eventually develop a uniform response to the same kind of particulars. But this does

[23] Salmon (1984), pp. 214–227.

[24] See van Fraassen (1980), Chap. 1.

not put us in a situation where we can conclude that this is the one and only way to group these particulars. We can say only that the capacity of consilience has been useful for the survival of those who had this capacity.

What is wrong with the realist's argument for the inference to the best explanation is not that no such inferences are used in science. But she needs to prove that it can be used to show that scientific success leads to referential success and that referential success cannot be explained away by the antirealist. Such an inference fails to prove that we are ontologically committed to those entities or laws of nature that are used by our best explanations. The argument works in favor of the realist's point of view only if she already has proved that we do have ontological commitments to the entities and properties postulated by theories that are empirically adequate.

4.4 The Success Argument

I propose that we distinguish between two sorts of scientific success. One kind is related to science's ability to conceptualize in a rigorous fashion the invisible world in terms of categories and principles, which allow us to make substantially correct predictions of numerous visible phenomena. Let us call this *theoretical success*. The other is related to our technological conquests of the invisible world and our ability to manipulate it to create new effects. This kind can be called *practical success*.

Theoretical success amounts to the fact that science until now has been able to foresee phenomena that no one could have imagined before their discovery. Science progresses into new areas, scientists build on already well-established truths, and it continuously accumulates further evidence of phenomena, some of which were hitherto unknown. Our best scientific theories have passed many empirical tests without being refuted, and they can yield coherent explanations of many otherwise unconnected phenomena. Therefore, it seems justified, the argument concludes, to consider those invisible entities postulated by a successful theory as real because they can be used to account for a large number of visible phenomena. For example, a concept like 'field' enters into a theoretical explanation of gravitational and electromagnetic phenomena; thus, the realist believes that we have every reason to assume that this concept refers to an objective feature of reality. However, if the invisible entity in question has been introduced solely for the benefit of a certain rather specific calculation, it seems less reasonable to believe that the term by which it is introduced refers to anything in the world, unless, of course, it helps the scientist to predict a new phenomenon as in the case of Planck's quantified oscillators.

Practical success makes science successful in virtue of our ability to construct an advanced technology on the basis of the insight into nature that we gain from applying scientific theories to practical problems. However, even though science by and large can be said to be successful in both of the above senses, the fact that science can be ascribed theoretical success hardly counts as a strong argument for scientific realism.[25] Theoretical success should be taken merely as evidence that current scientific theories are what they are supposed to be, namely, empirically adequate. In this case, explanatory success depends entirely on predictive success. Apparently, a scientific theory in physics cannot have explanatory success without having predictive success. But does it hold the other way around?

Sometimes it is argued that predictive success does not imply explanatory success as, for instance, in the case of quantum mechanics. It is held to be an example of a theory with very little explanatory power but with a lot of predictive force. Obviously, in this case the denial of the converse implication happens to rest on premises that are very sensitive to the kind of notion of explanation to which one subscribes.[26] However, with respect to the present discussion of what can be inferred from the success of scientific theories, it is not useful to make a distinction between predictive and explanatory success.

In the history of science, and even in contemporary science, there are many examples of theories that may be used to predict future phenomena, theories that are inconsistent, or make false empirical predictions, or whose central terms we come to believe do not refer to something real—e.g., the Ptolemaic system for the motion of the planets and Newton's

[25] Several philosophers share the view that theoretical success implies scientific realism. See, for instance, Boyd (1973), (1985) and (1990); Newton-Smith (1978) and (1981); and Niiniluoto (1977).

[26] For a further treatment of this issue, one may consult Faye (2014).

theory of gravitation.[27] In principle, the Ptolemaic model could still be used for predicting the observed positions of the planets on the vault of heaven, in spite of the fact that scientists no longer believe that the planets are satellites moving around the earth. Such predictions can be more easily achieved today because of the calculative power of current computers. Indeed, the geocentric model cannot explain all facts expected, e.g., the difference in movements between the inner plants and the outer planets, and this is one reason that today we think that the heliocentric model is correct. But in spite of the predictive success of the Ptolemaic model, nothing in reality corresponds to 'epicycles' and 'geocentric orbits,' the most central terms within the theory. Analogously, the world cannot be as Newton's theory of gravitation held if Einstein's general theory of relativity gives us the correct description on a much grander scale. The central term of the theory, "gravitational force," does not refer to something in reality; instead it has been replaced with "geodesic curvature of spacetime". However, the Newtonian theory is in fact still used for the prediction of many astronomical phenomena and solutions of technological problems in connection with space research, tidal movements, and virtually all terrestrial phenomena in our daily experience.

The conclusion is therefore that predictive success implies neither truth nor the reality of the presumed referents of theoretical terms. Scientific realism cannot make capital out of the fact that science has strong predictive success. The best that predictive success can prove is that the world works as if there were these entities. That some theories have useful predictive power without being true or having referential success may instead be interpreted as a general indication of theories more prescriptive nature.

But what about the converse implication: do truth and referential success imply predictive success? As Larry Laudan argues, scientific theories may be genuinely referential without being successful.[28] The examples he mentions are Dalton's theory of atoms, the Proutian theory that the atoms of heavy elements are made up of hydrogen atoms, and Bohr's

[27] It is well known that Bohr's model of the hydrogen atom contained internal inconsistencies. Nevertheless, Bohr was able to predict accurately the ionized helium spectrum and Sommerfeld the fine structure of the hydrogen spectrum. But, as Vickers (2012) concludes, this example seems to be a case against the realist's thesis that (theoretical) success is indicative of truth.

[28] Laudan (1982), p. 223.

early model of the electron. All of these were apparently genuinely referring theories in spite of fact that they made many flawed and inconsistent claims about atoms and their constituents and today are regarded as unsuccessful. Laudan also rejects a possible realist retreat, according to which it is said that a theory whose central terms refer will usually be successful. He does so because, as he says, it is always possible by the use of negation to generate "indefinitely many unsuccessful theories, all of whose substantive terms are genuinely referring." And he compares this logical point with the many unsuccessful theories of atoms that have been proposed during the two millennia of speculations about the nature of matter. If Laudan were correct, it would imply that the realist's argument at this point is badly damaged.

Nevertheless, I do not think that Laudan gives the realist sufficient benefit of the doubt. I believe that a realist with perfect justice may claim that various historical theories were not successful because some of their central terms did not designate anything or described it in a way that we no longer believe was correct. Some of them did, of course, since scientists had correctly identified those entities in question. But Laudan seems to imagine that the realist position asserts only that substantive terms are interpreted as referring. Against this, the realist could argue that the most important predicative terms also have to be genuinely referential for a theory to be successful. For example, a sentence like "Electrons move around the nucleus in energetically stationary, but otherwise classical orbits" expresses one of the fundamental assumptions Bohr made. Here the realist could argue that the terms "electron" and "nucleus" refer, whereas predicates like "move around in stationary but classical orbits" and "have a determinate position and a determinate momentum" were not satisfied. For this reason, while Bohr's theory made many impressively successful predictions, today we regard it as incorrect. It ascribed the wrong attributes to the right entities. So what made some of the theories mentioned unsuccessful was in fact that some property terms of the theories failed to be satisfactorily defined or turned out to be empty.

The above example also reveals how truth and reference are related for the scientific realist. Usually, the truth of a theory is taken to imply the genuine reference of its theoretical terms, while genuine reference does not necessarily imply the truth of very statement the theory implies.

A theory can only be true or approximately true if its terms have some real counterparts. In other words, whereas truth is, even according to realists, assumed to be sufficient for successful reference, reference is merely supposed to be necessary for truth. This is not the place to take a more careful look at the realist notion of truth, something that will be discussed in the succeeding chapter. But we still have to finish our discussion of whether scientific success is a parasite on genuine reference.

In addition to the putatively theoretical success of explanation and prediction, science is connected with practical and technical success. Maybe successful predictions are not a consequence of the fulfillment of the referential aspect of the theoretical terms employed, assuming that all that observation can justify is the genuine reference of the observational terms and hence empirically successful theories. Nevertheless, in science we are able to perform experiments with things that we cannot see with the naked eye. Afterwards, on the basis of the knowledge of the causal properties of these unseen objects that we obtain from these experiments, we may use them in the design of the physical function of technical apparatus and instruments. The realist could argue that because we can manipulate what we cannot see and bring about the observable effects we want to produce, we are justified in concluding that the theoretical terms of both the causal description of the experiment and of the function of the involved apparatus genuinely refer. It is an undeniable fact that we repeatedly, with increasing success, create and construct new technologies by using invisible entities to create causal mechanisms allowing these technologies to operate as we anticipate. But this fact would be inexplicable unless our best current theories were genuinely referential. For instance, since we are able to move around individual genes in a cell, taking some out and putting some others in, and thereby creating new organisms, it would be beyond any rational warrant to suggest that genes are not real merely because we cannot see them.

As Ian Hacking points out, the fact that electrons can be used as tools is an example of the strongest kind of evidence for realism about invisible entities.[29] But this is not necessarily an argument for the truth of these theories, i.e., for them to provide a literally true description of how things

[29] Hacking (1983), Chap. 16.

ultimately are. We do not need theories to know whether such entities exist or not. In his opinion it is not because one can make experiments with them that one is committed to believing in their existence, nor is it because electrons can be used to experiment on something else. What matters is that by understanding the causal properties of electrons, we can use our knowledge to build devices in which the electrons will behave in a certain characteristic manner whenever we want them to do so. Based upon some rules of thumb, electrons can be prepared in such a way that they can be employed to produce phenomena we wish to investigate in some other domain of nature.

For the entity realist, this conclusion is equivalent to holding that practical success implies referential success. Although the converse entailment is not true, theoretical terms may indeed have reference without the referent being an entity that can be used technologically. The unstated premise of the entity realist's inference is that you may "observe" something that does not exist and wrongly believe things are real that you cannot see, but you can never manipulate anything that is not there. And even less can you manipulate an entity to cause a wanted effect unless it exists. So the conclusion of the entity realist is that any theory of knowledge that confines knowledge to be about those things we can see with the naked eye is not very convincing.

A fine example illustrating some of these points is the discovery of hafnium.[30] The periodic system of the elements was not established until around 1870. When this happened, it was done only on the basis of the *chemical features* of the elements, and most chemists regarded it as a purely empirical classification of the elements. In 1897 J.J. Thomson suggested a connection between atomic structure and the periodic system, but it was not until Niels Bohr's second model of the atom that anybody was able to give a physically satisfactory account of all the elements from hydrogen to uranium, including the transition groups and the rare earths. The theory Bohr used to construct this account was a result of a mixture of ill-defined general principles and empirically based concepts coupled with an exceptional measure of physical intuition. Among the principles and

[30] My knowledge about the discovery of Hafnium rests entirely on an excellent study by Kragh (1979) and Kragh (1980).

theoretical concepts were the construction principle (Aufbauprinzip), the correspondence principle, penetrating orbits, and symmetry concepts. On the empirical side was chemical evidence in the form of ionic colors, magnetic properties, ionization potentials, atomic volumes, polarizability, and physical evidence in the form of optical spectra. Relying on these data and formative principles, Bohr gave a physical description of the atomic structure of the various elements and of how the electrons build up in shells from one element to the next. This description was able to reproduce many of the characteristics of the original periodic system.

After the formulation of Bohr's model, it was soon strongly supported by its ability to incorporate evidence from X-ray spectroscopy discovered by Dirk Coster. This evidence was in agreement with the predictions that included the right number of curves for the absorption edges, indicating the possible configuration based on levels defined in terms of three quantum numbers; the curves of absorption edges showed that the building up of electrons started out roughly where it was expected; finally, the curves almost reproduced those parts of the periodic system in which the atomic construction occurs at the intermediate, but still incomplete level. Likewise, the model predicted new results for optical spectra of the elements subsequently confirmed by Friedrich Paschen and Ralph H. Fowler.

Nevertheless, Bohr's model was overthrown a few years later, partly because J.D. Main Smith and E.C. Stoker changed it in order to cope with the structure and the existence of simple chemical compounds and partly because the exclusion principle introduced by Wolfgang Pauli could support their changes by explaining the electron distribution in each shell of a single atom. In spite of this achievement, Bohr's model still had another big victory to claim. At this time the element with atomic number 72 had not yet been satisfactorily identified. It was generally believed to be an element that belonged to the rare earths, and chemists were looking for it in ytterbium minerals. In 1911, Georges Urbain claimed to have isolated this new element by the method of fractionations. He called it "celtium." Eleven years later, Urbain, together with the X-ray spectroscopist Dauvillier, announced that, based on a few X-ray lines, they finally had identified element 72 in agreement with Urbain's earlier chemical discovery. However, had this claim been correct, it would have been fatal for Bohr's theory, according to which element 72 should be considered to be a homolog of zirconium and therefore have no chemical similarities with the rare earths as celtium was alleged to have. Knowing this and unhappy with the quality of Urbain's and Dauvillier's X-ray lines, Dirk Coster and George von Hevesy succeeded within half a year in finding the missing element, now named "hafnium," among zirconium minerals. They also used X-ray spectroscopy to identify the new element on the basis of two excellent lines that Coster showed were part of its L-spectrum.

When one focuses only on the predictive success of Bohr's model, as would van Fraassen, one could argue that it merely provided us with an empirically adequate account of the correlations of the various optical spectra of the elements and of the various X-ray spectra as well as a similar account of the mutual correlations between these two kinds of spectra. But the same does not hold with respect to Coster and Hevesy's manipulative success. Manipulation proves that invisible entities exist; it does not only show that our beliefs about them are empirically adequate. Why is this? Because we cannot think otherwise. We have evolved through natural selection to believe that invisible entities are real, just as well as visible, in case we can get in touch with them. Selection and adaptation have made systematically bodily interaction with the world our innate criteria of mind-independent existence. If we can intervene and influence things in virtue of our behavior we are automatically justified in believing that these things are separated from us. Epistemic skepticism is not a real option. Any reservation about invisible entities is not an actual possibility. Descartes' doubt about the external world is the imagination that plays its game upon the mind. Given that the human species are adapted such that our faith in external things is unavoidably associated with our bodily action, we are biologically perforce to accept the reality of things we can manipulate. It is simply imperative for our survival.

However, what makes this justification "automatic"? It is certainly true that we have a strong psychological tendency to believe this, and when we are in our everyday state of mind, we naturally do believe it, and the philosophical scruples of the scientific antirealist seem silly. So psychologically this belief seems "automatic," but that's not the same as logical justification, which is what traditional philosophy has demanded. However, rationality is not above psychology; it is part of psychology. Surely, rational thinking rests on certain norms based on reflection. As humans eventually gained the capacity of reflective understanding, the need for conscious standards of justification became immanent. But if these standards were in conflict with our innate cognitive mechanism, we would be much better off without any standards.

I think Hume would agree that *psychologically* antirealism is not a real option, but also insist that manipulation is *the only* justification. But why not just stop with highly confirmed models? Surely if we were to ask a physicist why he believes that positrons, say, are real, he would say because our best models, which tell us that these bits of visible evidence are the effects of invisible positrons, have been extensively tested and are well confirmed. It seems very unlikely he would say because we can spray them on niobium balls to neutralize their charge. So if we are going to rest the case on our psychological tendencies and dispositions, we do not always have to go to the level of manipulation. Evolutionary epistemology is not supposed to justify only manipulative success and ignore the clear selective advantage of organisms holding beliefs that exhibit predictive success. For example, predicting agriculturally relevant phenomena would have surely been of use to evolving *Homo sapiens*.

4.5 Constructive Empiricism

A theory of elements is empirically adequate if the world observationally appears as if there are such elements. Van Fraassen distinguishes between acceptance of a scientific theory and the belief in its (partial) truth, claiming that acceptance involves only the view that the theory saves the phenomena, not that it is true.[31] Nevertheless, the acceptance of a theory about *S* means to take all its claims literally, both claims about "observable" and "unobservable" entities. His idea is that by acceptance we commit ourselves to using the entire potential of the theory as if *S* exists in giving explanation and doing research. Still, we should be agnostic about the claims a theory makes about "unobservable" entities because they cannot be observed. Consequently, according to van Fraassen, the confirmation of Bohr's model would not force us to embrace a belief in

[31] van Fraassen (1980), p. 8 and p. 12.

the reality of atoms. The model was accepted for a while simply because it was considered to be empirically adequate in virtue of yielding successful predictions.

But is it possible to account for the discovery of hafnium without believing that Bohr's model of periodic system is true regarding its assumption about atoms? In more general terms: is it possible to accept a model without being externally committed to the theoretical entities it is a model about? That Coster and Hevesy were able to isolate and produce hafnium in quantities so large that everybody directly could see the material seems to justify a belief in hafnium atoms. As scientists eventually accepted the reorganization of the periodic system on atomic ideas, they began to accept ways to identify the different elements on the atomic level, which, I hold, at the same time established the referent of hafnium, even before samples of this element became visible to their eyes.

In my opinion we can accommodate invisible but "observable" entities into the ontology of common sense. As I see it, there is no epistemological difference between seeing a red car on the street and the instrumental observation of a cell. Both of the acquired beliefs are acquired non-inferentially as soon as one has a robust conceptual grasp of one's own visual experience. I assume that this ability of having non-inferentially determined perception was gained long ago in evolutionary history, and it is something we share with many animals. So, according to the evolutionary naturalist, the world of science is not merely the empirical world of perception. The premises tacitly employed in seeing the objects of common sense are hard wired by natural selection, but the theoretical principles involved in observing or detecting the invisible entities, which enlightened common-sense realists believe exist, are neither common to all humans nor hard wired by natural selection. They are based on fallible assumptions of human-created science. Take an object like gamma ray bursters. We have arrived at the view that they are enormous explosions in very distant galaxies, but that is a conclusion reached by a lot of research and controversy. Astrophysicists think of the issue as settled today, but it seems totally possible that future models, or empirical discoveries, might replace the ones accepted today and our belief about what causes gamma ray bursts will be altered. Then we will say our belief that GRBs are caused by enormous distant explosions turned out to be false;

these "objects" do not really exist. I do not think we could discard the hard-wired premises of common-sense realism like this, or if we tried to, we would soon perish; while solipsism is a philosophical possibility, it is not a real possibility for organisms evolved through natural selection.

Elsewhere I have argued for a criterial theory of meaning according to which the observational criteria for identifying each element are part of the meaning of the name of that natural kind.[32] There is a causal connection between the use of the name and its bearer. The causal connection is determined by the criteria we have elected to use to identify the bearer of the name; in the present case of hafnium, the evidence was in the form of chemical data, particular lines in the optical spectra and in the X-ray spectra. These evidential empirical criteria are satisfied by the bearer's sortal properties, and they enter into the definition of a particular name "hafnium" and determine the reference of that name.

The mere fact, however, that Coster and Hevesy could manufacture a new visible element by extracting invisible atoms hidden inside zirconium minerals seems unintelligible if we only think of the periodic system as an empirically adequate classification. The last point can be stated even more dramatically. A couple of elements between hydrogen and uranium do not occur in nature as, for instance, technetium. It is a metallic element that can be obtained by bombarding molybdenum with deuterons or neutrons. Now, if the only thing you do is to change one visible element into another visible element by adding invisible things to it, are you not justified in believing that these invisible things exist?

When microphysical processes can be deliberately manipulated in a purposeful and constructive manner, then we have strong and justified reasons to assume that our belief in the existence of atoms, protons, and neutrons is true. It would also be impossible to explain the success of our technological innovations unless we were able to refer to microphysical entities and to tell an approximately true causal story about them. We are able to do this since scientists understand their causal properties, and therefore they can use that knowledge in performing experiments and measurements. In general, technological success requires that beliefs about what we are doing with invisible entities have to be true,

[32] Faye (2002), pp. 72–78.

and these beliefs can be true only if we are capable of identifying the entities involved and have knowledge about their causal behavior.

In order to explain why the use of imperceptible entities and magnitudes implies beliefs, and not merely acceptance of the theory, as van Fraassen suggests, Sam Mitchell has developed a functional argument for why it has to be so.[33] First, he lays down a condition that should be acceptable for an empiricist like van Fraassen: Only if somebody would act differently toward two kinds of entities does it make sense to argue that he or she harbors different kinds of epistemic attitudes toward these entities, that is, holding a certain belief is true on one hand and being agnostic toward it on the other. Then he argues that in designing experiments and measurement apparatus, scientists treat perceptibles and imperceptibles in no discernibly different way. Therefore, van Fraassen must claim either that we should be agnostic about the existence of perceptibles, too, or that we should believe in the existence of imperceptibles, too. But since van Fraassen seeks to found our attitudes toward imperceptibles on our justification for accepting them (namely that claims about them are part of an empirically adequate theory) rather than in believing in the approximate truth of the theory we accept, then the justification for believing in the "observables" of the theory should be sufficient for believing in the so-called "unobservables" of the theory. Moreover, I contend that there are no obvious epistemic grounds on which to draw a demarcation between visible and invisible entities. And thus the distinction will not bear the weight van Fraassen puts on it.[34]

The criterial theory of meaning, according to which the causal relationship between the name and its bearer results from the criteria for identifying the bearer, allows the scientist the possibility of changing these criteria. The use of a "natural kind term" is always open to revision because the criteria are revisable for some pragmatic purpose. Whenever science discovers that what is regarded as identifying criteria does not refer to sortal properties, we may delete some of these criteria and replace them with new ones, or we may enlarge the number of remaining criteria, or in the worst case

[33] See, Mitchell (1988).

[34] See, for instance, Faye (2000).

scenario, we may give up the idea that a certain set of criteria establishes reference to a genuine entity as it happened with caloric, phlogiston, etc.

Van Fraassen's constructive empiricism confines our ontological commitments to those physical things we can see with our naked eye, and in that sense he celebrates *plain* common-sense realism. The evolutionary naturalist extends his commitments to invisible entities that can be observed by instruments, and in that sense he supports *sophisticated* common-sense realism. However, both are skeptical about the representational view of scientific theories.

4.6 Structural Realism

No doubt, the realist has a strong case if he refers to the technological applications of science as something that is sufficient to explain the referential success of scientific models. The practical success of science supports the external ontological commitments of the language of science. It is because of their manipulability that invisible entities are acceptable to sophisticated common-sense realism. Notice, furthermore, the difference at this point between theoretical success and practical success: it is only the latter that is sufficient for referential success. Technological progress is a result of our power to act and intervene in physical processes. It shows that there are parts of the world that we cannot immediately see with our unaided eyes but to which we have cognitive access through instrumental observations. But, taking this for granted, it still remains to be proved that this kind of progress could not be explained on the assumption that the manipulated reality always exists as a conceptually grasped set of entities, properties, and relations and that perhaps these might be described in another way if the cognitive abilities of human beings had been different.

The kind of realism we have opposed takes the present scientific theories to be true or approximately true about the things in themselves. Because of the optimistic no-miracle argument, it holds that only true theories can explain the success of science. Laudan has, in contrast, introduced the pessimistic meta-induction argument: The existence of theory change in the past seems to supply good inductive grounds for holding that presently accepted theories sooner or later will be replaced by new theories. Therefore, predictive success guarantees neither truth nor reference. Laudan's argument assumes that reference cannot be secured externally independently of the truth of the theory that first introduced them. The physical content of a theory permits it to be true or false, but then if a theory eventually is overturned by a new one, truth cannot be what explains the empirical success of a theory. In the attempt to stay clear of this dilemma, some realists argue instead that theories have empirical success because of the structure of mathematical formulation of a theory. This view, which John Worrall attributes to Henri Poincaré, but which he was first to explicate, is called structural or syntactic realism.[35] This form of realism, he argues, can account for no miracle argument and meets Laudan's objection that scientific realism is unable to explain the transition from an older theory to a newer one where the latter is inconsistent with the former. Structural realism gives us the best of both worlds and still explains why succeeding theories have empirical success.

Structural realism is not a full-blown realism. The idea is that science may completely misidentify the nature of things as they are described by the metaphysical and physical content of our best theories but still attribute the correct mathematical structure to reality. Worrall says, "The rule on the history of physics seems to be that, whenever a theory replaces a predecessor, which has however itself enjoyed genuine predictive success, the 'correspondence principle' applies."[36] This requires retention of formal structure over the change of theory in the sense that the mathematical equations of the old theory reappear as limiting cases of the mathematical equations of the new theory. Worrall's historical case is the transition from Fresnel's to Maxwell's theory of light. Fresnel's theory made correct predictions because it accurately identified certain relations between optical phenomena that depended upon something undergoing periodic motion at right angles to the light.

But what, more specifically, is a structural realist? It cannot be that a realist interpretation of the meaning of scientific theories yields an understanding of the physical content of the laws of nature. In his discussion of this problem James Ladyman mentions that structural realism

[35] Worrall (1989), p. 112.
[36] Worrall (1989), p. 120.

may be formulated as one of two alternative positions: an epistemological refinement and a metaphysical approach.[37] The epistemic structural realist holds that there are epistemic constraints on what we can know about the world. We are justified in believing that we possess objective knowledge if there happens to be a mathematical continuity across theory change and scientific revolutions. This proposal requires a clear-cut distinction between the structure and the content of our theories, that is, a distinction between the mathematical equations and the theoretical interpretation of the formalism.

It is possible to find some support for this view in Bohr's methodology for developing a theory of quantum mechanics. Bohr introduced the principle of correspondence, and no other physicist has made such an explicit use of the correspondence principle as a guiding principle in the formation of a new theory. He realized that according to his model of the hydrogen atom, the frequencies of radiation due to the electron's transition between stationary states with large quantum numbers, i.e., states far from the ground state, coincide approximately with the frequencies to be expected from classical electrodynamics for a free electron. However, his own model of the atom eventually failed to predict some of the spectroscopic phenomena that were observed in the years to come, and by the beginning of the 1920s it was quite obvious to Bohr and other leading physicists that they still had to look for the final theory. Hence, in the search for a consistent mathematical formalism that could predict all observations, it became a methodological requirement for Bohr that any further theory of the atom should predict values in domains of large quantum numbers that should be a close approximation to the values predicted by classical physics. The correspondence rule was a heuristic principle meant to make sure that in areas where the influence of Planck's constant could be neglected, the numerical values predicted by such a theory should approach those predicted by classical radiation theory.

The correspondence rule was an important methodological principle in the historical development of quantum mechanics. In the beginning it had a clear technical meaning to Bohr: It demanded that calculations based on the mathematical formalism of classical electrodynamics gave

[37] Ladyman (1998), p. 410.

the same result the ones got by applying the rules of Bohr's "old quantum theory" as one approached the limit of large quantum numbers. Using the correspondence principle, Bohr was able to connect the frequencies of radiation of an atomic spectrum with the Fourier components of the motion of an electron in orbit and then "compare the radiation emitted during the transition between two stationary states with the radiation which would be emitted by a harmonically oscillating electron on the basis of electrodynamics."[38] Later on, Bohr as well as Heisenberg considered quantum mechanics as a mathematical "generalization" of classical mechanics in which certain structural elements are preserved.[39] Matrix mechanics fulfilled the promise of the correspondence principle in retaining the forms of the classical equations in the limit where the quantum of action could be ignored.[40] Accordingly, we can explain the predictive success of classical physics if we take into account that it agrees with quantum mechanics in the domain where the effects of the quantum of action play no significant role.

In contrast to modern structural realists, however, Bohr realized at the time he became involved in the interpretation of quantum mechanics that in order to get to the meaning of quantum mechanics it was not sufficient to preserve only some structural features. The formalism cannot be understood unless we continue to interpret it in such a way that classical concepts must be used to describe experimental results, and we therefore have to apply these concepts while interpreting the mathematical formalism.[41] I think Bohr was right. It is obvious that it makes no

[38] Bohr ([1920]1976), p. 51.

[39] Although Pauli's exclusion rule and the introduction of spin broke with the attempt of explaining the structure of the basic elements along the lines of the correspondence argument, as Pauli pointed out in a letter to Bohr, Bohr continued to think of it as an important methodological principle in the attempt to establish a coherent quantum theory. In fact, he repeatedly expressed his opinion that Heisenberg's matrix mechanics came to light under the guidance of this very principle. In his Faraday Lectures from 1932, for instance, Bohr emphasizes: "A fundamental step towards the establishing of a proper *quantum mechanics* was taken in 1925 by Heisenberg who showed how to replace the ordinary kinematical concepts, in the spirit of the correspondence argument, by symbols referring to the elementary processes and the probability of their occurrence" (1998, p. 48). See also Heisenberg (1967), p. 98.

[40] Bohr ([1925]1984), p. 852.

[41] Faye (1991), pp. 113–119, presents a formal semantic formulation of Bohr's principle of correspondence.

sense to compare the numerical values of the quantum mechanical predictions with those of classical physics unless the meaning of the physical terms in both theories is somehow commensurable. So in Bohr's opinion the use of the correspondence principle in developing the new quantum mechanics justified the epistemic ideal that classical concepts, like position, momentum, and energy, are indispensable for our understanding of physical reality at the empirical level, and only when classical phenomena and quantum phenomena are described in terms of the same classical concepts does it make sense to compare the predictive results of different mathematical formalisms. Therefore, this example shows that the structural realists' attempt to draw a revealing metaphysical distinction between structure and content, i.e., between formalism and interpretation, fails. Worrall's structural realism is unable to explain the predictive success of theories because it focuses on mathematical structure separated from its interpretation in the context of a physical theory. One can develop pure structures in a priori mathematics, but when one employs some bit of these mathematical structures to express a physical theory (with empirical content) then the "interpretation" issue becomes crucial. To explain predictive success requires attributing to the phenomena not only structural, but also substantive properties as well.

Ladyman also rejects the epistemological form of structural realism; he claims it offers no advantage over traditional scientific realism. His objection considers two possible ways of understanding scientific theories. One way is to look at a theory as a Ramsay structure in the sense that a Ramsey sentence for the theory replaces the conjunction of all theoretical constants with distinct observational variables bound by existential quantifiers. The result is that theoretical terms are eliminated but the observational consequences are preserved. It is a mistake, however, to think that this Ramsification of a theory entirely eliminates the theoretical terms. They are still being referred to, not directly with theoretical terms, but indirectly via their Ramsey descriptions, whose direct referents are known by acquaintance. The idea is here that the world consists of "unobservable" entities between which "observable" properties and relations obtain. Thus, the relations form the structure of the world, the structure itself is the abstract form of a set of relations that hold between these entities, and the relations are those that can be known. William Demopolous and Michael Friedman have shown that the problem with this linguistic structural understanding is that any structure of a set of relations can obtain from any (sufficiently large) collection of objects. But if that is the case, no given structure picks out a unique set of relations characterizing this world. Therefore, we should reject a Ramseyian understanding of the structure of a theory.[42]

An alternative conception has been proposed by Stathis Psillos, a reading that makes structural realism indistinguishable from traditional realism.[43] He argues that Worrall's mathematical continuity is not sufficient to answer the pessimistic meta-induction; we need a positive argument connecting the mathematical formalism to predictive success, an argument that shows that mathematical formalism successfully represents the structure of the world. He also doubts that it is possible to discriminate between our ability to know this structure and our ability to know that it successfully represents the nature of the world. Instead, he argues that structure and nature are inseparable; properties are defined by the laws in which they feature, and the nature of something consists in its basic properties and their relations as they are structurally described in mathematical equations.

Ladyman advocates an ontic or metaphysical version of structural realism because he claims that this is the only option that can explain ontological discontinuity. He (and Ross) defines his position in these words:

> Ontic Structural Realism (OSR) is the view that the world has an objective modal structure that is ontologically fundamental, in the sense of not supervening on the intrinsic properties of a set of individuals. According to OSR, even the identity and individuality of objects depends on the relational structure of the world. Hence, a first approximation 'There are no things. Structure is all there is.'[44]

What we normally take to be individual entities are nothing but structures all the way down. There are relations but no relata, at least not

[42] See Newman (2004) for a criticism of Ramsey sentence realism posed by Cruse & Papineau (2002).

[43] Psillos (1995) and Psillos (1996).

[44] Ladyman and Ross (2007), p. 130.

considered as self-subsisting individuals existing in space and time. One consequence of this view is that there are modal or nomological structures, but these structures have no causal power or efficacy. What we think of as causal structures exist only as pragmatic approaches in the special sciences. Compared to modern physics, there are objects, but they are deprived of any inherent nature, identity, and individuality.

The ontological commitment of structural realism is more than to the empirical content of a theory but less than to the full ontology of scientific realism. Ladyman also thinks that the ontic approach to mathematical structures fares well with the semantic or model theoretic view of theories because "theories are to be thought of as presenting structures or models that may be used to represent systems, rather than as partially-interpreted axiomatic systems."[45] The predictive success of specific theories of physics, such as starlight being bent near the sun as predicted by general relativity, is possible to understand if we assume that the most abstract mathematical structures go beyond a correct description of actual phenomena and represent modal relations between them. He opts for an elaboration of structural realism that takes "structure to be primitive and ontologically subsistent."[46] He then draws attention to Weyl's view on objectivity, according to which the status of objectivity can be bestowed only on relations that are invariant under particular transformations. Thus, ontic structural realism takes structures and relations to be the fundamental constituents of the world rather than objects and properties.

According to Ladyman and Ross, ontic structural realism "as we develop it is in principle friendly to a naturalized version of Platonism …"[47] So how do they relate to Sellar's manifest image? This is a figment of evolution. Here is a couple of quotations: "To say that all that there is are relations and no relata, is therefore to follow Plato and say that the world of appearances is illusory,"[48] and "Individual things are locally focused abstractions from modal structure."[49] Moreover, the manifest image cannot explain the

[45] Ladyman (1998), p. 416.
[46] Ladyman (1998), p. 420.
[47] Ladyman & Ross (2007), p. 158.
[48] Ibid., p. 152.
[49] Ibid., p. 153.

use of representations in fundamental physics. In contrast, scientific realists mistakenly believe "that the appearances are caused by unseen objects and that the behaviour of these objects can be invoked to explain the appearances. But the resources of the manifest image cannot be (directly) used for satisfactory representation in physics. Hence, mathematics has an ineliminable role to play in theories."[50]

Ladymand and Ross distinguish between structures in the formal sense, which they simply call structures, and structures in a material sense, what they call "real patterns." The former are the mathematical models by which we map the real patterns that make up the relations between empirical data. Of all possible structures that could possibly obtain, how do we pick out the "right" one? It would seem that logically there must be many that would be equally consistent with any finite set of data points. This is also a question that occupies van Fraassen: How is it possible to distinguish between non-instantiated mathematical structures from instantiated physical structures?[51] The calamity is that Ladyman and Ross have no satisfactory answer to what that difference is:

> Physical structures exist, but what is it?…What makes the structure physical and not mathematical? That is a question that we refuse to answer. In our view, there is nothing more to be said about this that doesn't amount to empty words and venture…The "world structure" just is and exists independently of us and we represent it mathematico-physically via our theories.[52]

Such a confession cannot hide that this in-the-world Platonism cannot explain why, say, some geometrical structures, supported by empirical evidence, seem to represent the world structures, while others do not. A conventionalist could object that the choice of geometry depends on the conventional selection of a metric and not on a factual matter, but since they are also representationalists, ontic structural realists cannot embrace such a possibility. Einstein's General Theory of Relativity connects spacetime structure to the distribution of matter. Although I shall later discuss

[50] Ibid., p. 158.
[51] Fraassen (2006).
[52] Ladyman & Ross (2007), p. 158.

our understanding of the space-time structure, Einstein's idea relies on a common-sense conception that a structure and what is structured are two sides of the same coin.

Likewise it is difficult, at least for me, to apprehend how physical information can be recorded by physical instruments if the recordings take place as interactions with pure modal structures. How can pure modal structures provide us with information at all? How can that which is not physical "interact" with that which is purely physical? The word "interact" is well defined in science only for physical interactions; that is why talk about the brain and the mind as "interacting" is uninformative. The exchange of energy, for instance, cannot be characterized in terms of mathematical structures due to the fact that we cannot determine the "energy" of a mathematical structure. In modern physics, energy and mass are two forms of the same thing, which we call "matter." Even though the exchange of energy between a system and an instrument can be seen as a registration of information about some verifiable property of the system, and therefore qua verifiable can be considered as a relation, it is also well known that quantum mechanics itself offers no description of the measurement process. It describes only the development of the system in probabilistic terms. [53]

Some philosophers have raised objections to the ontic version of structural realism, but I do not intend to present these in any detail.[54] I want to say only that among realist-minded philosophers the difficulties of turning a mathematical representation into an ontological description of structure have been acknowledged. For instance, Mauro Dorato and Frederico Laudisa nicely summarize Maudlin's view that "mathematical representations of physical phenomena are not a clear guide to ontology, since they often do not guarantee even isomorphic relations between themselves and the latter. Furthermore, for obvious algorithmic reasons they must greatly simplify and idealize the target they are a vehicle for, and so they are not

---

[53] Of course that would be a counterargument only if one also assumes QM is "complete."

[54] See, for instance, Pooley (2005): "The main thesis of this paper is that, whatever the interpretative difficulties of generally covariant space-time physics are, they do not support or suggest structural realism" (p. 2). Also Friggs (2002) delivers an extensive criticism of structural realism. He convincingly argues that representation and structural isomorphism are not enough for embracing structural realism because most physical systems can be represented by different structures. What is needed is the human intention to establish a representation for a particular purpose.

necessarily similar to what they are supposed to denote."[55] Maudlin himself points out that "mathematical objects acquire algebraic and numerical properties that the physical objects do not have; there are purely gauge degrees of freedom in the mathematics."[56] Ontic structural realism may attempt to overcome these problems by claiming that physical phenomena are nothing but mathematical structures represented by mathematical equations. But such a move is no way out. How can the concrete manifest world existing in space and time "emerge" from a causally inert world not existing in space and time?

Apart from what I have already said, my own disagreement with Ladyman's ontic view rests on the following considerations: First, the semantic view of theories assumed by Ladyman is not necessarily an advantage for the structural interpretation. Not all proponents of the semantic theory of theories consider themselves realists. Bas van Fraassen provides one example. Moreover, the semantic view of theories is beset with some of the same problems as is structural realism. Both rely on assumptions that are difficult to defend. On the one hand, the immediate interpretation of a theory is taken to be a model of abstract objects; on the other hand, a theory consists of a set of descriptive sentences, each of which has a certain truth value.[57] According to an ontic structural realist, who focuses on structure rather than content, physical theories represent a concrete structure, which means that a scientific theory is true or false with respect to some concrete relations and structures in nature. However, if we understand the meaning of a mathematical equation in virtue of knowledge of abstract relations, how can we determine the truth of the equation expressing a physical law in virtue of the actually existing structures of this world? The structure of a theory does not correspond directly to some real structure but to the structure of some models, which constitute the interpretation of the theory, i.e., a set of mathematical expressions is structurally coherent with its models, and one of them

[55] Dorato & Laudisa (2014).

[56] Maudlin (2013), pp. 152–153.

[57] Cf. Faye (2006) for further criticism of the semantic view on theories. See also Faye (2002), Chap. 8. Instead of considering a model as part of the interpretation of a theory, I take it to be a construction to which a theory applies in virtue of defining a vocabulary and a set of rules using the vocabulary.

may then be isomorphic with the real structure of the world. It remains a puzzle to me how we can understand a particular theory's structure by having access to the abstract structure of the models.

Second, ontic structural realism seems to represent a naïve view of the relationship between mathematics and reality made familiar to philosophers from Wittgenstein's old picture theory of language in *The Tractatus*. The ontic structuralists see mathematics first and foremost as a means of representing the world in thought. The function of mathematical formulas is to represent how the world is structured. This is possible only in so far as the meaning of a mathematical equation is interpreted in virtue of a corresponding structure, which, if it is realized, makes the mathematical formula as so interpreted true. As Wittgenstein argued with respect to language, any combination of sentences consists of a relation of logical structures of atomic sentences, and these atomic sentences stand in a direct relation to the corresponding possible facts so that the sentences are isomorphic with the atomic states of affairs they picture. Similarly, a mathematical formula forms a structure itself, and this structure is "interpreted" by saying that the world is structured in the same way as the formula in order for it to become true. In this sense, the mathematical structures are logical pictures of possible real structures. The mathematical structure of an approximately true physical theory mirrors or pictures approximately the structure of factual relations. Thus, our currently best scientific theories and reality exhibit a mutual isomorphism by having the same structural form.

Setting aside the later Wittgenstein's criticism of the picture theory, there is, I think, an important difference between his attempts to grasp the function of language in terms of the atomic sentences that picture possible facts and the ontic structural realists' attempts to understand the function of scientific theories in terms of mathematical structures that are isomorphic with some possible factual structures. Wittgenstein's idea was combined with the assumption that we have direct empirical access to the facts that were pictured by a language, say, the cup is on the table. But structural realists cannot have a similar empirical knowledge of the *modal* relations of the world since these relations are ontologically independent of the "emergent" entities that take part in them. The object of theories is mathematical structures, real counterparts to our mathematical equations, but we have no plausible way to get to know their existence by direct empirical inquiry, i.e., the modal laws that are the physical counterparts to the equations are inferred from the repeated behavior of like individuals. All we can observe and manipulate are objects and their properties.

Third, it does not suffice for the structural realist to begin with the ontological commitments to structures given to us by accepting certain theories. The commitment to a certain structure is always internal to the mathematical framework. The ontic structural realist needs to point to some external commitments in order to move the structure from merely a logical possibility to that which is instantiated in reality. Again, I think that Bohr pointed to some fundamental problems concerning the mathematical structure of our current physical theories to the effect that no such external commitments can be made for the equations expressing many of the principles of fundamental physics. In both quantum mechanics and relativity theory we meet complex numbers in the formulation of some of the basic questions such as the commutation rule and the four-interval invariant relation. Therefore, Bohr rejected the view that theories give a 'pictorial' representation of the world.[58] His reasons seem to be that mathematical structures, which appear as a result of the use of imaginary numbers, can never be the object of our experience and therefore cannot be known to be instantiated in the world in itself. The existence of imaginary numbers is clearly a human construction by means of mathematical abstraction from real numbers. This releases us from having any external commitments with respect to the structure of such theories.

Fourth, representations need not resemble what they are supposed to represent. In those cases where the representans can be compared with the representantum we may sometimes realize that little resemblance exists between those two. For instance—and here I need to consider only a more earthly example than those of fundamental physics—if we attempt to represent the upsurge of carbon dioxide in the atmosphere as causing the increase of the world's average temperature in terms of logical relations, it is not very plausible to argue that a logical representation such as $\forall x \forall y (P(x) \Rightarrow Q(y))$ or $(\sim p \;\Box\!\!\rightarrow\; \sim q)$ is structurally similar to the causal relation. Is the material or the counterfactual implication more

[58] See Bohr (1999) p. 86 and p. 105.

structurally similar to causation than the other one? Whenever we see a resemblance between representans and representandum, say between a picture and the pictured, it is because human beings have chosen which observational features should be mapped by the representation. When we express a causal relationship in a logical equation, we do not observe the causal nexus but only the temporal succession of events. The structural realist infers from the structure of language to the nature of the world. But is this inference legitimate? I do not think so. The reason for rejecting the inference is that the structure itself does not have any *explanatory* role to play. It is not the structure of the logical expression we use to describe the causal relationship that explains why the addition of carbon dioxide to the atmosphere produces a rise in the global mean temperature, rather the structure of the language helps us to make certain deductive inferences or calculations from one expression to another.

The final objection I briefly want to present is this. Assume that scientific representations are in general empirically underdetermined. Models may therefore be empirically equivalent without having the same content or structure. The mere fact that in principle it is possible to construct such models with different content or structure should make us suspicious of the ontological claims of structural realism. If the same observable facts can be described satisfactorily by structurally different models, we have no reason to argue that mathematical equations represent objective relations and therefore no empirical grounds to prefer one particular formulation rather than another.

Thus, it is clear that ontic structural realism manifests an indefensible position on the relationship between mathematically formulated theories and the world. An isomorphic coherence exists between the mathematical structures, which exist independently of the world, and the real structure of the world as it exists independently of mathematics. This assumption makes sense only if both mathematics and the world are designed according to the same principle of reason that allows a "mapping" of the logical relations between the elements of the world into logical relations between mathematical elements. In this way, a universal logic functioning as a superior principle for both mathematics and the world guarantees epistemological objectivity. This is all fairly mystical. In contrast, I believe that a less speculative and more practicable approach to an understanding of mathematically formulated theories and their relations to the world does not go via syntax and formal semantics, but through an approach to science that may involve ideas from cognitive semantics.[59]

4.7 The Failure of Representationalism

Behind both theory realism and structural realism lies a metaphysical presupposition that our best scientific theories represent either the world's fundamental entities or its basic structures. Indeed, in both science and everyday life we use representations, but from an evolutionary and pragmatic point of view, little support can be given to such a representational view of scientific theories. In general, representationalism is dear to what Dewey famously criticized as the "spectator theory of knowledge." In its place pragmatists wanted to place the knower not in the position of spectator, but in the position of interactor with nature. I read the failure of representationalism as a fundamental pragmatist lesson, which leads to the view of theories as conceptual tools.

An evolutionary naturalist holds that perception should not be regarded as a passive *representation* of the external world but an active *presentation* of it. Through perception and interaction, the external world displays itself to us by providing us with experiential information of the kind we have been naturally selected by evolution to receive. The conceptualization of sensory information is part of this adaptation. The consequence is that whenever we see something as a particular type of thing it is constitutive for our non-inferential perception. It does not rest on any interpretation.

Interpretation and representation go hand in hand.[60] Both are results of an intentional act of understanding in terms of some preferred conventions. What represents is a purposeful construction of what is represented. Maps are fine examples of representational devises. Here we can distinguish between the representans and the representandum and compare them according to the role of the representans. A map functions as a map always with respect to a definite purpose, which is that its structure helps

[59] A preliminary attempt along these lines can be found in Giere (1988) and Giere (1999). For a criticism of his semantic view on theories, see Faye (2006).

[60] See Faye (2014), Chap. 4.

to orient us in the world. The map has been constructed with the purpose of displaying some structural similarities to what is being mapped, similarities whose recognition in the map as well as in reality has to be visually accessible in order for the map to fulfill its purpose. But scientific theories are not like maps. A map works as a representation because it can be compared visually with the representantum, but a theory cannot be compared visually to anything other than other theories. This is one of the reasons why I take theories to be conceptual tools, expressed as linguistic rules, for the construction of models.[61]

If 'representation' always contains an intentional feature, which I think it does, one may be inclined to claim that no matter what is suggested to be a representans, it cannot but represent what it is intended to represent. However, the purpose as such is a necessary but not a sufficient condition. I think it is quite obvious that we need to put some empirical constraints on any understanding of representation. The purpose is not simply the intention of representing, but the intention of being able to compare the representans with the representandum in certain ways. If we cannot empirically demonstrate that they are similar to one another with respect to a chosen convention, we merely possess unredeemed high hopes that our alleged representation represents something. Thus, the lack of empirical comparability affects the premises of metaphysical realism, scientific realism, and structural realism as well.

In my opinion, it makes sense to talk about true theories or approximately true theories only if one believes that scientific theories represent something. I do not see how one can avoid being a representationalist concerning theories and at the same time claim to be a theory realist or a structural realist. Instead, I hold that scientific theories, like natural languages, are not representing anything but are tools for communication of our beliefs; nevertheless, you can still be a sophisticated common-sense realist by being committed to invisible entities but not to true or approximately true theories.

At this point, we might face the challenge of why successful theories are so successful and try to defend truth in the correspondence sense; presumably, the presence of successful theories is the reason why many realists find themselves committed to correspondence. However, the pragmatist rejects

[61] Ibid.

this as a dead end and defines "truth" in terms of the process by which we have arrived at it as the most "successful." In a sense the fundamental task of pragmatism is to define what "success" is and the method by which such successful theories are developed. Personally, I take the success to be associated with their adequacy of being conceptual tools for constructing models in terms of which the scientists can explain the world.

Where does this criticism lead us with respect to scientific theories? Invisible entities exist. But we do not need scientific theories to be true, or approximately true, in order to discover the existence of invisible entities. Invisible entities can be, and often are, discovered without scientists having any "approximately true" theory at their disposal. We are committed to their existence whenever we are able to interact with them in a constructive way. The truth of scientific theories is not needed because the relation between theory and entities is mediated by models. The entities such as planets, stones, pendulums, light, atoms, electrons, photons, and quarks are not, and will not be, the direct objects of any theory.

I have elsewhere argued for a non-representational view of theories. My own view is that scientific theories are neither true nor false. This is not traditional instrumentalism because I do not think that the claim implies that invisible entities do not exist. Theories are conceptual tools expressed as defining rules of interpreted formal languages. Fundamental laws, like Newton's laws, Maxwell's laws, and Schrödinger's equation, function as definitions by stating relations between set of quantities.[62] A theory consists of a vocabulary of certain idealized properties, which are defined as variables in some mathematical equations. The equations interrelate quantitative terms by defining some of them in terms of the others. Not until a mathematical model is established, which is an idealized representation of some concrete objects, will these defined quantities be part of descriptive statements that intentionally refer to the properties of concrete entities. We can then use this idealized model to explain the behavior of the corresponding physical entities. The upshot is that since past and present theories of physics do not deal with concrete entities but only mathematically define attributes, scientific theories may change without affecting our ontological commitment of the entities involved.

[62] See Faye (2014) pp. 103 ff. Also Faye (2002) Chap. 8; and Faye (2005).

5

Truth, Language, and Objectivity

We are realists by nature. Believing in an external world has added positively to the survival of our distant ancestors. But we are not born with only a natural sense of an external world; we are also born with a natural sense of truth: It would not have helped our long gone predecessors much to survive if they had had a sense of an external world but had no sense that their imaginary thoughts could give them a wrong representation of it. We benefit biologically from a genetically determined conviction that some beliefs correspond with how things are, whereas other beliefs may fail to do so. Even an innate sense of an external world and a similar one of truth would not have been of any advantage for our forefathers had they not also acquired a cognitive machinery by which they could establish which beliefs were in fact true and which not. The means they acquired were mechanisms to establish the fitness of thoughts to sensory experience. Ultimately, from a biological perspective it is what we can see, hear, and touch that causally determines what is true. It was first much later in our evolution that our reflection began speculating, reaching the conclusion that the external world may consist of things we could not experience.

Perhaps we are still missing something of importance; maybe things are not as simple as they look from an evolutionary point of view. Ontological

J. Faye, *Experience and Beyond*,
DOI 10.1007/978-3-319-31077-0_5

commitments, according to some modern metaphysicians, are concerned with truth-makers in contrast to Quine's view that these commitments are concerned with what entities the bound variables of the theory must range over for the theory to be true. As Ross Cameron puts it, "the ontological commitments of a sentence are not what the sentence quantifies over but rather what entities must be included in our ontology to ground the truth of the sentence—what entities must exist to make the sentence true."[1] This also holds for any sentence, which implies that its truth-maker transcends empirical accessibility. Metaphysical realism violates the Kantian limits, which deny the possibility of scientific knowledge of a thing-in-itself by claiming that facts about these are what ground the truth-makers. For a metaphysical realist the aim of any metaphysical inquiry is what really makes propositions truely uncontaminated by any human cognitive contribution. Thus, common-sense realism seems to involve one notion of truth; metaphysical realism relies on quite another. Common-sense realism is the view that measures truth with respect to how external things appear to us. What is true is true in an epistemic sense and not necessarily in a metaphysical sense. A true belief is what is best to believe, and that is what can be justified by empirical evidence and good reasons. But the metaphysical realist is not satisfied with this restriction of the application of truth. To him or her a belief is true regardless of the existence of any empirical evidence we may have for its acceptability. The problem for this person is how can we ever know when these conditions for truth-making are fulfilled?

The concept of truth is traditionally deeply ingrained in any metaphysical point of view concerning how reality really is, not merely regarding how reality is conceived by us. This reality-in-itself is regarded as existing independently of any cognitive beings and as inaccessible to direct experience. Therefore, this reality can be assumed to be very different from the world of common sense; however, metaphysical realists believe that there is more to say about reality than what even science can ever possibly determine. Science has been able to supersede the world of common sense and formulate abstract theories that can be used to explain things as these are experienced by us, but the ultimate reality, as the metaphysical realist understands it, may exist independently of the empirical objects

[1] Cameron (2010), p. 252.

of science. In his bones the metaphysical realist believes that truth goes far beyond what common-sense experience and theoretical science can inform us about. There are truths that are absolutely true independently of human experience. Neither common sense nor science in general can justify the truth of its own presuppositions and therefore show why everyday knowledge or scientific knowledge is possible.[2] The world is necessarily what it is, and truth depends solely on reality itself. Truth arises whenever we get things right about reality, even if we are unable to know by the use of our senses, or instruments, that what we believe to be true is in fact true. At least some metaphysical realists believe that we may discover these truth-makers by an a priori inquiry into the domain of reason alone. Realists might defend their view by arguing that if it does not make sense to believe in absolute truth; that is, if somehow truth has to be conditioned and constrained according to our cognitive faculties, then we seem to be in the absurd position of claiming that we know certain truths without knowing what makes our beliefs true, i.e., nothing independently determines a belief to be true. As soon as the realist embraces the existence of such self-subsisting truth-makers, we can have beliefs about the world-in-itself that are absolutely true or false. Although we cannot observe these ultimate entities or structures, realists believe we may infer their existence by reason and thereby be justified in our beliefs about them.

So the notion of truth that metaphysical realism requires seems to be the view that our beliefs about the world are true or false absolutely, because these beliefs stand in a particular relation of "correspondence" to the *real* world's truth-makers that endows them with a truth value regardless of any possible observational justification. The real world, we believe, is not merely separated from our immediate experience and imagination. The real world is also thought by rational reflection to be as it is untouched by our cognitive capacities, and therefore it could in principle be quite different from what our perceptions or observations take it to be.

In contrast, the metaphysical antirealist claims that the world is as it appears to us. He holds with respect to the separation of the perceptual world from our actual perception that we are confirmed in believing in the world of experience every second of the day. This is the reason why

[2] See Lowe (2002), pp. 6–7.

we are sometimes mistaken. But he does not know what the real world is apart from the empirical world. What is true is what we can establish to be true according to our senses. However, I think there are really two levels of antirealism here. One is agnostic about an independent reality; all we know is the empirical reality of the world as we see it. The other maintains a positive doctrine about this so-called "external" world. The second version holds both that truth goes beyond our immediate experience and yet also claims that what can act as a truth-maker is nonetheless species-dependent. A metaphysical antirealist, subscribing to evolutionary epistemology, would then argue that we are biologically adapted to experience the perceptual world as existing independently of ourselves as an "external" world, although the conscious awareness we have of this external world separated from our actual perception is a result of reflection. So those, and only those, thoughts, beliefs, or sentences that reflect the world as it is observed to be can be said to have a truth-maker. But we are not adapted to hold—and the metaphysical antirealist will reject it—that there exists such a real world, which makes any belief absolutely true or false. Instead he argues that the notion of a world as it is in itself is a mental construction.

I shall call metaphysical antirealism grounded in evolutionary epistemology "evolutionary naturalism." The position does not deny the existence of truth-makers. It constrains its scope of inquiry into their nature of existence as it is revealed by the empirical sciences, and it uses only those conceptual resources that evolution can possibly have given us to understand them. So an evolutionary naturalist would deny that there are truth-makers behind what we can observe by claiming that such a transcendental notion of absolute truth arises because we have misused our faculties of reflection to suggest the existence of something beyond our experience that these faculties are not adapted to represent as it is in itself. What interests us here is therefore the question of the intelligibility of the metaphysical claim that each and every true proposition is always made true in virtue of the real (not merely mistakenly imagined) existence of a truth-maker. Is it a misuse of our faculties of imagination and reflection to have a concept of truth that completely surpasses our cognitive abilities to settle whether or not it holds? In a nutshell, the problem is that if evolution has not furnished us with any cognitive means of establishing how the constructive elements of our reflective thinking, such as truth-making grounded in things-in-themselves, may or may not correspond to reality, it seems not to make much sense to claim that truth as correspondence with reality is the case for absolutely all propositions irrespective of its constructive nature. Assuming it does not make sense, our task is to develop an alternative naturalized notion of truth, which is less metaphysically ambitious, though more in accordance with what we know about human evolution and our realist instincts.

Here I shall opt for a notion of truth that is true to our realist instincts that not only is there an external world that makes our beliefs true, but also that a correspondence relation between a truth-maker and a belief can only be justified by pointing to the epistemic means that help us to establish the correspondence. I shall argue that the truth relation is not primarily an internal relation supervening on its relata, but an external relation constructed in terms of our conceptual comprehension of what counts as truth-makers. As Ruse succinctly puts it: "Obviously, working within the common-sense level, the Darwinian is just as much of a correspondence thinker as anyone else…But at the final level, defending common-sense reality, as we have had to accept, the Darwinian subscribes to a coherence theory of truth, believing that the best you can do is to get everything to hang together."[3] Usually these two approaches to truth are considered to be in opposition. That they need not be is still something we have to demonstrate. I attempt to bring them together in a three-stage model of truth.

5.1 What Is Truth?

We are interested in truth because we want our beliefs to be correct; we want them to represent the world as accurately as possible. In everyday life we use the term "truth" to tell each other that what is said is as it is said to be. This holds in science too. The claim of the scientific realist is that our best theories are true, or approximately so, and that truth as such is a goal worth pursuing in making theories, simply because we are thereby told that the world is as our theories represent it to be. The

[3] Ruse ([1986]1998), p. 202.

desired connection between theories and reality gives us truth. A sentence about an objective state of affairs is not always true, but only so in cases where the sentence expressing a belief is somehow compatible with the way things are. Therefore, metaphysical realists see truth as a feature bestowed on our opinions, beliefs, and sentences in virtue of a correspondence between them and the reality they are representing. Such a correspondence is assumed to exist objectively, perhaps even unnoticeable by the person who harbors a belief or asserts a declarative sentence. The metaphysical realist also holds that a belief or a statement acquires its truth value immediately with the formulation of this very belief or statement, quite independently of whether or not we have means to establish its relation to the facts. The truth relation is considered to be *internal* in the sense that its existence does not rely on our epistemic capacities: it is enough for the relation to hold that the relata exist; that is, if the possibility that would make a belief true is actually the case, when the belief is factual, then when the belief is formed, it becomes automatically true.

Thus, the traditional understanding of the notion of metaphysical truth welds together an ontological component and a semantic component by means of the notion of correspondence: A belief or a sentence is true if, and only if, it represents a fact or a state of affairs, irrespective of whether or not this fact is cognitively accessible. For instance, experience never reveals what is necessarily the case or possibly the case, but only what is the case. The correspondence theory of truth catches our common intuitions that the truth of a certain belief, thought, or declarative sentence is determined by certain facts independently of that very belief, thought, or declarative sentence and that these facts are represented by the belief or the sentence in an appropriate way. Facts are the truth-makers and beliefs, thoughts, and sentences are the truth-bearers. But before this linkage is warranted, we must know what truth is and how it can be related to facts and meaning.

On the one hand, truth seems to be a feature of particular thoughts, opinions, judgments, beliefs, and assumption and, on the other hand, a feature of particular sentences, remarks, utterances, or statements expressing these thoughts, beliefs, etc. But either way these tokens can possess the feature of being true only because they represent some facts, i.e., they have a semantic content corresponding to the facts. So apparently we can discern between, say, the thought itself and what the thought is about. The former is, one might say, a psychological state of somebody's thinking, and that state is not less real than tables and electrons are, whereas the latter concerns the semantic content of his or her thinking. Thoughts, opinions, beliefs, or sentences have content because they are considered to function as mental (or physical) representations of objective states of affairs. It is this function, and this alone, that conveys thoughts, opinions, beliefs, and sentences with the possibility of having a truth value. The capacity of our thoughts (and statements expressing these thoughts) to represent something other than themselves is what we have in mind when saying that they are true or false by correspondence. For that reason alone one might expect 'truth' to be in the eyes of the beholders of these thoughts and beliefs. If this is so, the metaphysical notion of truth as correspondence would prove to be deeply incoherent.

Whatever theory of truth one wishes to promote as the most cogent view, the notion of truth intuitively contains certain self-evident features constraining all reasonable theories. Crispin Wright mentions seven basic principles that a theory of the truth predicate must satisfy, although he writes them in a different order: (1) a content is true just in case it corresponds to the facts; (2) truth and warrant are distinct; (3) truth is absolute, i.e., something cannot be more or less true; (4) truth is stable, i.e., if a content is ever true, it always is; (5) truth is explicable by Tarski's disquotational schema; (6) an assertion is to present something as being true; (7) every assertible content corresponds to an assertoric negation.[4] In the subsequent account we shall relate our discussion to most of these principles to see how far they can be incorporated as axioms into a naturalist theory of truth. As we shall see in due course, especially (3) and (4) can be put into question even by metaphysical realists.

Let me begin by saying that I take belief tokens or sentence tokens, like statements and utterances expressing those beliefs, to be the main bearers of truth. Only particular beliefs and particular sentences can represent the world as it actually is. Belief types or sentence types can be true only by definitions, i.e., not in virtue of their representational role, since every representation is context-dependent and only tokens appear in any

[4] Wright (1992), p. 72.

context. So, if I am correct, truth is a property that belongs to particular beliefs and statements of these beliefs. This move avoids problems with the truth of token-reflexive sentences whose content varies with the context of utterance. Some philosophers, however, regard truth as a property of propositions. As Paul Horwich says, "No doubt we do attribute 'truth' to statements, beliefs, suppositions, and so on; but surely what we have in mind is that the propositional objects of these linguistic and mental acts are true, and not the acts themselves."[5] But propositions are commonly considered as abstract entities by expressing (or intending) the informational content of what is believed, stated or supposed such as "that the Earth is spherical." So if we attribute truth to a proposition, we are ontologically committed in advance to the reality of an abstract realm of meaning existing independently of any context.[6] Besides making a mystery out of how we get to know this world, the theory fails to explain why the same token-reflexive sentence, being stated in different contexts and by different persons, apparently expresses the same proposition although it may have different truth values in these different contexts.

Instead, we shall say that a proposition can be expressed by a true or false sentence, and a sentence is true or false whenever it is stated or uttered, i.e., the sentence gains a truth value as soon as it is formulated rather than after the accumulation of evidence by the person who formulates it. The semantic content of a complete declarative statement is the kind of state of affairs the statement describes. Those states of affairs are believed, stated, or supposed. Moreover, I hold that facts, if there are such things, are actualized states of affairs (thus, states of affairs are possible facts). Indeed, the meaning of the sentence determines what it can be used to describe (since words conventionally signify those things they stand for), and it therefore determines the semantic content of the sentence. But whatever supplies a sentence with meaning, the sentence does not necessarily have a content detached from the state of affairs it describes.

[5] Horwich (1990), p. 17.

[6] In contrast to Horwich, Searle (1978) argues that the truth conditions of all sentences reflect some background beliefs about the context in which it is appropriate to utter the sentences.

A metaphysical notion of truth consists of two other elements that we shall dismantle: (1) the world consists of objective facts whose nature may be different from how they appear to us; (2) beliefs or sentences are true in case they stand in an appropriate relation to these objective facts. If a statement is true in virtue of a correspondence to some facts, "correspond" as well as "fact" must be well understood, elucidating the kind of relation and its relata in order for them to take part in an explanation of truth. Needless to say, one of the weaknesses of the notion of correspondence is recognized to be a non-metaphorical specification of what it means for a statement to correspond with objective facts. It does not make us much wiser by being told that the statement 'mirrors' or 'pictures' the facts, nor is it clear how we can identify facts without referring to those judgments by which we assert these facts. Indeed, nothing is wrong a priori with a view that truth is equivalent to corresponds-to-the-facts. But the fact, if it is a fact that the predicate "is true" might be substituted with "corresponds with the facts," does not engender a theory of truth that explains what it means for a belief or a sentence to be true. Whether the notion of correspondence turns out to be relevant or not for a theory of truth depends entirely on whether it is possible to give a proper, non-circular clarification of this notion of corresponding to facts that enables us to elucidate our intuitions about truth.

One of the adherents of the correspondence theory of truth, Bertrand Russell, saw the corresponding relation as a "rich" relation of congruence between language and reality. He believed that correspondence consists of a structural similarity between the constituents of propositions and the constituents of facts. Today it is generally agreed that it is incomprehensible how such an isomorphism between the meanings of those words that make up a sentence and the physical things, properties, and relations that constitute a fact is alleged to hold. Ignoring this is, I believe, the cardinal sin that structural realism makes. How can a proposition, "The cat is on the mat," in any nonconventional way represents the fact that the cat is on the mat, if the cat, the mat, and the position relation of being-on are the physical constituents of the fact? A similarity of structure between the truth bearer and the truth-maker is neither necessary nor sufficient for truth. Putnam is right when saying, "everything is similar to everything else in infinitely many respects."[7] One must therefore be able to specify in which ways the representans and the representandum are alike before the alleged isomorphism can be recognized. But such a specification requires that one has access to the representans and to representandum independently of each other. This is possible only with facts within the range of human perception. But the problem becomes acute when talking about things that are defined in a way to make them beyond experience. If we are referring to empirical facts, then it seems always to be possible just to point to the fact as long as we are only asserting true statements. To deny that an empirical fact is the case we cannot point to a non-fact. Neither does it make sense to say that a statement is a collection of physical signs, pictures, physical things, and properties, for a picture may perfectly match an object without being a picture of that object, and it may not resemble an object and still be a picture of this object.

Already Frege—in spite of seeing himself as a Platonic realist—argued most cogently against the idea of defining truth in terms of correspondence to the facts. However, parts of his objection were already contained in the German idealists' set of arguments in favor of truth as coherence, a view they inherited from Kant. The basic complaint is this: it makes sense to say that something is an accurate picture in case it corresponds with what it depicts because we can compare the picture with the object it is supposed to reproduce, that is, because we can link the picture to reality. But it makes no sense to say that a statement, or the thought it expresses, corresponds with the state of affairs when we talk about truth. Here it is not a question of discovering a fact and then comparing it with the statement. A declarative statement cannot be compared with the non-interpreted state of affairs it intends to describe in order to see whether it represents it correctly or not.

Another version of the correspondence theory of truth is not troubled by similar pitfalls. This version is due to John Austin who saw correspondence as a correlation between language and reality. According to him, such a correlation rests entirely on linguistic conventions that are not intended to mirror or picture anything at all. Convention reflects our intentions. He operated with two kinds of conventions: *descriptive*

[7] Putnam (1981), p. 64.

conventions correlating types of sentences with types of situations, things, events in the world and *demonstrative conventions* correlating statements with historic, or actual, situations in the world. A statement is then said to be "true" when a demonstrative convention correlates it with the historic state of affairs of a type that is descriptively correlated with the sentence used to state it.[8] In other words, Austin apparently held that a correspondence theory of truth has to be based on the notion of reference: a statement is true if, and only if, it is refers to the appropriate historic situation, and a situation can be said to be appropriate for making the statement true if, and only if, it is of the type which meets the conventional requirements of application of the sentence with which the statement is expressed. In this proposal Austin took some reasonable assumptions for granted: First, a statement always refers to a particular state of affairs; second, a declarative sentence of a certain type is by definition descriptively correlated with a certain type of state of affairs.

Austin's theory is attractive, but it contains some loose ends and unsolved problems. First, there is some ambiguity in connection with the central concept of descriptive and demonstrative convention. It is unproblematic, I think, when Austin holds that the descriptive correlation between a type of sentence and a type of situation guarantees the conventionality of using tokens of this sentence to refer to a situation of this type. Such a prescription determines that the statement that "*X* is *F"* is correctly carried out with no other sentence than the sentence "*X* is *F*." But what makes Austin's claim problematic is that the relation between the statement that "*X* is *F"* and a state of affairs in the world, the relation that we assert obtains when we utter the sentence that "*X* is *F*," is purely conventional and therefore changeable, if we would like it.[9] As Geoffrey J. Warnock truly points out, it is a question of fact, not of convention, whether or not a statement is true. He also notes that Austin's definition of demonstrative convention is such that the convention correlates a certain statement with a particular situation. But he adds that this is inconsistent with Austin's own conception of a 'statement,' according to which a certain statement is partly identified by the aid of the situation to

[8] Austin (1950), p. 22.

[9] See Strawson (1950), p. 44, who is very critical of Austin's view, and Warnock (1962), p. 67n.

which it refers. Thus, in order to avoid such an inconsistency I think one must say something like this: It is the speaker's recognition of a particular situation as belonging to a certain type that makes the speaker utter this particular sentence on this particular occasion. That he chooses this particular sentence rather than any other possible sentence for his statement is because he see the actual situation to be one in which the sentence will be the correct one to state (because he recognizes the situation as a kind with which it is correlated). However, the type of sentence of which the statement is a token determines what kind of facts the statement concerns. A sentence type and the type of a state of affairs partly help to identify each other.

But this solution inevitably raises another question about the individuation of facts, for the same fact may be identified by different descriptions. The fact that Roald Amundsen was the first man on the South Pole can also be referred to as the Norwegian with the biggest nose. The right answer is, so it seems, that a statement can be identified as a physical sentence token being uttered in time and space independently of the particular state of affairs, and the actual state of affairs can be identified as a physical phenomenon in space and time independently of the actual statement. Hence we can use different descriptions to identify the same state of affairs.

Second, Austin's theory can also be criticized for want of the ability to explain the nature of states of affairs with which negative, existential, general, hypothetical, disjunctive, and counterfactual statements are correlated. The paradigmatic case of his theory is the singular, affirmative statement like "The cat is on the mat." But what about declarative statements like "The cat is not on the mat," "There are white cats," "All cats are mammals," or "If the cat had not been on the mat, it would have been in the garden." If these statements are true or false, does it mean that the first statement is demonstratively correlated with the particular negative fact, as Russell suggested; the second with the particular fact of some cat's being-or-not-being white; the third with the particular fact of all cats' being-or-not-being mammals; and the fourth with the particular counterfactual fact of the cat's being-or-not-being on the mat and would-be-or-would-not-be in the garden? It is not quite obvious to believe that such facts exist that we cannot see and know how to recognize. For if we cannot experience such particular facts, it is difficult to see that they should instantiate types of situations to which the sentences "Some cat's …" and "All cats' …" are descriptive correlated. The problem does not improve when we look on putative modal, causal, hypothetical, counterfactual, subjunctive, and probabilistic facts, or tensed facts for that matter. Granting for the moment that singular and positive facts exist, why not go the whole hog and claim that these other kinds of facts also exist? Nothing stops one from accepting all kinds of facts if one first accepts some of them. Such facts must be real if facts make such statements true or false. So the question is whether there are facts that correspond to these different types of sentences like those that allegedly correspond to singular, affirmative empirical statements, but from which we are barred from having any empirical access, and so do not have any perceptual possibility to decide whether exist or not.

After having raised the issue, we can either accept that facts may exist regardless of whether we can (empirically) recognize them or not, as the metaphysical realist would do, or we can doubt, as the evolutionary naturalist does, that they exist at all. To the evolutionary naturalist there are no facts that are beyond the empirical facts that we can observe are the case. It is, of course, tempting to insist on the view that there are some states of affairs with which we are directly acquainted, say, as when we see a book lying on the table. All those things, persons, and properties we are familiar with being related to one another seem to be the naked facts we are looking for. Nobody in his senses, except the skeptic, would reject his experience of a book on the table when he sees one there. The book on the table leaps to the eyes—it is simply a paradigm case of witnessing a fact. If we are not acquainted with books on tables, we are not acquainted with any facts at all. But we are: It is incoherent to say both that there is a book on the table, and it cannot be a fact that there is a book on the table. Things and events have various properties and stand in various relations to one another, just as they lack other properties or fail to stand in various other relations to each other, and their lacking these properties or failing to stand in these relations are also facts. Through our sense experience we are continuously informed about such facts and the experience we acquire by our senses causes us to believe that there is a book on the table. Biologically speaking, we have no other options but to rely on the information we get through our senses.

But the issue of the general ontological status of facts is not the main issue; I assume both realist and antirealist admit that *empirical* facts are 'real' and that we are directly acquainted with them. Traditional antirealists stopped there. There are no facts other than empirical facts—what we can sense directly. Now scientific realists want to go further and say there are trans-empirical facts about unobservable entities that explain the observable facts; indeed, they cause these empirical facts. The problem with these facts is that our beliefs about them are the results of inferences from observable facts, which also use theoretical laws, assumptions, etc. Thus, their status is epistemically more precarious, so to speak. They might be facts, and actually we have good reason to think they are facts, but they are not as certain as what we can know directly. In contrast to both, I have argued that their views of what belongs to the empirical world are too narrow. I see no reason why one should limit the empirical world to the world we can directly perceive. An evolutionary naturalist may argue that invisible entities can be observed by instruments although they are not directly perceivable by our senses. So in her view atoms, molecules, etc., are part of the empirical world to which we are adapted even though we cannot directly perceive any of them. Still, the empirical world consists only of things that appear to us because atoms and molecules appear to us in the form of perceptual effects we consider as good empirical evidence for their existence. A trained scientist may immediately (non-inferentially) understand the observational evidence as the presence of this or that 'theoretical' entity because this evidence is part of his or her concept of being this or that entity. Such our grasp of instrumentally produced experience results from an inductive process of learning how to conceptualize this experience. For the evolutionary naturalist, facts about such invisible entities are not trans-empirical facts and therefore epistemically on par with facts about visible entities.

Thus, it seems natural to assume that what provides the correspondence between a statement and reality is the referential nature of representations. First and foremost we can say that a sentence corresponds to some state of affairs because various parts of the sentence conventionally refer to different parts of reality. Take, for instance, a simple affirmative sentence like "*X* is *F*." Here the sentence would be true in virtue of the existence of an object *X* that "*X*" represents and that belongs to the class of objects "*F*" applies to. The sentence can be true because it has a referential structure. Thus, the utterance "*X* is *F*" is true if, and only if, there exists an object *X* such that "*X*" refers to *X* and "having *F*" is satisfied by *X*. Facts are actual states of affairs referred to by sentence tokens in the way just mentioned, whereas all states of affairs are possible facts that are identified by the meaning of the corresponding sentence types. But how will this account handle the other kinds of states of affairs?

Austin argued, correctly I think, that negative facts à la Russell do not exist. A statement is not false because it corresponds with a negative fact existing in the real world but because it misrepresents positive facts. If a statement is true, it reports an actual fact; if it is false, it reports a possible state of affairs that is not actualized. Suppose "*X* is *F*" and "*X* is not *G*" are both true; this would be equivalent to saying that "*X* is not *F*" and "*X* is *G*" are both false. Now "*X* is not *G*" is true because the object to which "*X*" refers belongs to the class of objects that may not be members of the class denoted by "*G*" and *X* actually does not satisfy the criteria for being in *G*. Similarly, "*X* is *G*" is false because "*X*" refers to an object that is a member of the class of objects to which "*G*" may not apply and actually does not apply. But all states of affairs are stated by declarative sentences that may either be true or false. So why is it not possible to perceive negative facts? Because negative facts are not present in experience; they are inferred from what is present. Be that as it may, this particular response to the problem of negative facts does not yield an appropriate characterization of a fact, that is, a general response to answer the question of whether there exist existential, universal, nomological, causal, modal, counterfactual, or probabilistic facts. If facts are what make sentences true, these true-endowing entities must exist objectively in all their plenitude independently of the type of sentences that can be used to express them. But if facts exist objectively, they should be distinguishable by other means than by their linguistic counterpart.

Consequently, on the one hand, if a fact is merely a consequence of the claim that every well-formed declarative sentence token in a language is true or false, and then we are internally committed to believing in the existence of possible facts corresponding to every type of sentence. If all well-formed declarative sentences in a language have truth value by corresponding to facts, the existence of different types of sentences will be sufficient for the identification of different kinds of facts. But then the idea of facts making a sentence true seems to be vacuous, for a fact is just what can be expressed by a true statement. On the other hand, if a fact is something to which we are externally, or objectively, committed, it means that we must have a general notion of facts as real independently of any language with which they are taken to be descriptive correlated. Yet, such a view immediately gives us a problem. Granting that language-independent facts make every type of declarative sentence true or false, we face the problem of how to identify objective facts that transcend every possible means of verification, such as counterfactual, modal, and universal facts. Can we solve this dilemma one way or the other?

5.2 Truth and Meaning

The problem with the existence of language-independent facts is closely associated with the question of whether a theory of truth can be formulated independently of a theory of meaning or whether a theory of meaning can be stated independently of a theory of truth. Whether we can answer in the affirmative, or in the negative, the question persists: how do truth and meaning relate to one another? These questions are notoriously some of the more difficult ones in philosophy of language, because separate intuitions, which appear to be reasonable apart from each other, may be in mutual conflict with one another when put together.

Let us assume that the meaning of sentence types determines the kind of possible facts or states of affairs that may obtain. Meaning is capable of performing such a job because meaning determines the conditions under which any sentence token is true or false and because these truth conditions in turn determine the reference of the sentence. Such an assumption presupposes that we have some means for explaining meaning independently of a theory of truth. Consequently, we cannot identify any fact separately from the various features of sentences in a language. What can count as true depends entirely on the language itself, since 'truth conditions' become internal to the language, a notion that is determined by the way we classify the various sentences semantically. A language containing sentences of existential, universal, nomological, causal, modal, counterfactual, and probabilistic form will each be assigned a truth value according to its coherence with other declarative sentences. What can act as a fact is no longer a language-independent issue. The notion of correspondence between an external truth-maker and a linguistic truth-bearer as a feature of truth fails to play an independent function any longer.

Truth as correspondence can fail for two very distinct reasons. Either truth is really not a genuine predicate at all, or truth is a genuine predicate, but cannot be analyzed in terms of correspondence. The first view is presented by the so-called "redundancy" theory of truth. It holds that the predicate "is true" is empty because the ascription of truth to a statement is nothing more than what the statement says itself. The application of the truth predicate to a statement is simply tantamount to asserting the statement itself. We do not get more information about a statement by being told that it is true than we do just by hearing it be stated. The truth predicate is superfluous. Ramsey was the first who explicitly introduced this idea, but already Frege appeared to have a similar position.

The conclusion, which Frege drew from his opposition to the theory of correspondence, was quite evident: Truth cannot be a relation between propositions (i.e., sentences, beliefs) and facts. Instead facts are nothing but true thoughts, simply because there is no difference between saying "It is a fact that Napoleon married Josephine" and saying "The thought that Napoleon married Josephine is true." However, Frege's objections against the correspondence view of truth pull the rug from under only the Russellian version of isomorphism, whereas Austin's explanation of truth in terms of correspondence-cum-correlation evades Frege's objections. Austin specifically denied that "The fact that *X* is *F*" must lead us to believe that "fact" means the same as "true statement." The reason why we can say "He stated a true statement" is just because we believe that his statement refers to a fact that is believed to have been described correctly. The phrase "It is a fact that …," or its equivalent "It is true that …," is intended for being used in situations where we ignore the distinction between language and the world and speak simultaneously about them. But this does not settle the issue whether at least some facts are real and can be seen.

Apparently, the redundancy theory of truth is most famously formulated in Tarski's theory of truth. A minimum requirement for something to be true is generally considered to be expressed by the so-called equivalence thesis: (T) *s* is true in *L* if, and only if, *p*, where "*p*" stands for a sentence that is a reiteration of the object language sentence *s* or is a translation of the sentence "*s*" is standing for. Sometimes Tarski's disquotational theory is taken to be identical with the redundancy theory. The difference between them is, however, that the redundancy theory holds that "truth" and "assertibility" are synonymous, whereas the disquotational theory holds that 'truth' merely has the same extension as 'assertibility.'

The disquotational theory of truth captures some of our intuitions behind the basic principles mentioned in the beginning of the last section. The disquotational schema states that if a sentence is true, then the sentence is assertible, and vice versa. The redundancy theory, however, takes for granted that a sentence is true because it is assertible. It is the feature of 'assertibility' that explains the use of the truth predicate. Indeed, the opponent of the redundancy theory will not deny that the notion of assertibility is somehow logically connected to truth, but he will insist that the implication is the other way around: A sentence is assertible because it is true, or believed to be true. It is the truth of the sentence, or our beliefs about the sentence's truth value, which explains why it is asserted.

The coherence view holds that a sentence cannot "correspond" to anything other than itself. The assumption is that whenever we claim a sentence to be true or false, we do it on the basis of the sentences in the language we already assume to be true or false. We can only relate sentences to other sentences, and consequently a sentence can only be true or false if it coheres with the entire system of sentences; that is, if the sentence we want to evaluate is consistent with the coherent system of "true" sentences, it will also be true. This is due to the assumption that a speaker of a certain language understands, as Quine would say, a particular sentence only because she understands the whole language. The meaning of a sentence can be explained only in terms of the meanings of other sentences that again have to be explained in terms of the meanings of further sentences, and so on.[10]

It is not difficult to find some attractive aspects attached to this suggestion, which seems to reflect our linguistic practice as competent speakers. Assume somebody holds a certain sentence to be true, say, "There are five great tits on the feeder in my garden," and he is asked what makes him believe such an assertion is true, how would he then respond? Well, if he begins by saying that he believes that the sentence is true because the sentence corresponds with the fact, we will not think that he was really explaining to us his reason for asserting it as true. If we do not know anything about the existence of great tits, such an appeal to the fact will not convince us at all. What we want of him is an explanation of why he considers it to be a fact that there are five great tits on the feeder. The proper way to do so seems to be by making an appeal to other sentences that both of us hold to be true: "There are five birds on the feeder," "They have a black head, white cheeks, and a yellow crest divided by a black stripe," and "These birds have a size similar to a house sparrow." Subsequently, he can point out that according to the description of a great tit in Peterson's field guide, all these true sentences indicate that the sentence in question is true. Thus, by referring to other sentences that we both accept as true, he is able to tell us why the sentence is true. However, my point is that it is not only us, but also the speaker who gets an explanation of the reasons for ascribing a certain truth value to that particular sentence.

No matter how attractive the theory of coherence is and how much it accords with our linguistic practice, it also has some unpleasant implications. First, it holds that there is really no exit from the linguistic system to the world. Every explanation of truth is always in terms of truth of other sentences where we may end up in a circle in which such an explanation contains the original sentence. This is a serious problem in itself.

[10] Of course this argument predates Quine! But it just seems silly because many states of affairs can be simply pointed to or located by conventional indexicals that reach outside the circle of language. If this were true, "radical translation" of the sort Quine considers would be simply impossible; everyone would be totally a prisoner of his own language. There could be no "gavagai" scuttling through the bush, because in this strange world ostention is forbidden; we must keep our hands in our pockets.

But our common intuition about truth is that something in the world makes empirical statements true or false, and it is not the truth value of other sentences. It is, indeed, something about a language as a whole that helps making a particular sentence true, since the language has been constructed to describe and communicate what one can observe whenever one observes it. But a declarative sentence is not always true, its truth value changes as the world changes, and therefore as our observation changes. So it must be something apart from the meaning of the sentence itself that makes it true. In order to break out of the linguistic confinement, we need some ways of reaching out to the world that can explain our intuition that it is something in the world that acts as the truth-maker.

Second, the coherence theory seems to conflate the truth of a belief and its warrant. When a person explains her belief in the truth of a particular sentence by citing her reasons for the assertion of the sentence, she is not explaining why the sentence is true, but why she believes it to be true. Only if the meaning of the sentences in the explanans is analytically connected to the meaning of the sentence in the explanandum can the truth of the former act as evidence for the truth of the latter. Of course, the meaning of a sentence has to be such that the evidence for the sentence's truth and what makes it true cognitively have to coincide in order for us to have the ability to recognize its truth value. We have to be able to acknowledge the kind of truth relation that exists between the truth-maker and the truth-bearer based on the evidence available to us. If such a belief in the truth of the sentence is not simply a result of chance, the relationship between the truth-maker and the evidence for the truth-bearer must be stronger than a mere contingent relationship. It has to be of such a nature that whenever we have sufficient evidence of truth in, we then know that we have—at least for all practical purposes—identified the truth-maker.

A theory of meaning must give an account of what we understand when we grasp the meaning of a sentence in a language. Here the traditional wisdom is that we gain this understanding by knowing the conditions under which the sentence is true or false. The consequence is that a theory of meaning depends on a satisfactory theory of truth. For if the truth conditions of a sentence determine its meaning then understanding the truth conditions is a precondition for understanding meaning. Thus, there are at least three serious objections to semantic antirealism: (1) We have to know the meaning of a sentence before we can verify whether it is true or not. The rejoinder might be that the meaning of a sentence is not a question of whether the sentence token is actually true or not but is a question of what would prove it true or not must be cognitively accessible by humans. A reply to the rejoinder would then be that we still have to know the meaning of a sentence in order to know what would count as its verification. (2) There may be different proofs of one and the same statement as with the geometric and the arithmetic proof of Pythagoras' theorem. The only way to avoid this problem is to say that the meaning of a sentence is not identical with the procedure of proving its truth. (3) There are groups of meaningful sentences, for instance, about the future, where we are forever prevented from establishing their truth values. How can their meaning be spelled out in terms of assertibility conditions? As we shall see, the evolutionary naturalist has a view of truth and meaning different from the semantic realist and more similar to the semantic antirealist.

5.3 Non-Realism Concerning Truth

We have defined realism as the view that external things exist independently of the mind. A proponent of metaphysical realism claims that a statement (belief, etc.) is true only because a certain external state of affairs objectively obtains. A metaphysical realist theory of truth is therefore a theory that states that (1) the very same state of affairs, which the statement is about (belief in, etc.), obtains independently of the mind; (2) it is this objectively existing state of affairs that makes the statement true or false independently of the existence of any cognitive being's recognition of this state of affairs. In opposition to this metaphysical realist theory of truth, we have all non-realist theories. These can either deny the ontological component of the metaphysical realist theory that there exists an objectively mind-independent world or deny the semantic component that the facts of the matter are what determine truth.

On the one hand, theories rejecting the ontological component are traditionally associated with solipsism or with subjective or absolute idealism, whereas, on the other hand, theories rejecting the semantic element usually subscribe to the idea that truth is somehow an epistemic notion.[11] According to these theories, the truth relation has something to do with justification or warranted assertibility. In the previous chapters, I have already dealt with the ontological component. Therefore, here I shall consider the other kind of theories that focus on the epistemic notion of truth.

The metaphysical realist takes the relation between the truth-bearer and the truth-maker to be objective or somehow given independently of any relation to cognitive subjects. But what does this claim really mean? It seems obvious, indeed, that if somebody claims something determinate about particular things that can be experienced, say "The cat is on the mat," such a sentence token is true or false regardless of whether any person (or animal) perceives this state of affairs or is even capable of perceiving the cat on the mat. For, if the world is objective, individual facts do not change according to whether or not we witness them. And, since the sentence token and the fact are both objective, the truth relation between these relata has to be objective itself. Thus, the truth relation between the utterance and the fact of the matter exists irrespective of whether anybody is looking at that particular cat.

In this sense truth and warrant are distinct: we separate the truth relation from the grounds for judgments of truth, i.e., the grounds for asserting whether the fact makes the utterance true or false. The latter is based on perception or some other confirmation producing method. When it comes to individual statements (beliefs, etc.) concerning particular physical states of affairs, perception does not constitute the relation between the truth-bearers and truth-makers; instead, it establishes the reasons for ascribing a certain truth value to sentence tokens (belief tokens, etc.) by the assessment of facts. It is impossible to see how one could hold that particular states of affairs exist objectively and at the same time claim that these facts need to be perceived in order for them to make a sentence true or false. Which property does the perception add to an actual state of affairs that it did not have before, and what turns the state of affairs into a truth-maker? Any semantic non-realist theory of truth that does

[11] Kirkham (1992) contains one of the best and more recent expositions of different theories of truth including various epistemic-based theories.

not grant a separation between the relation of truth and the relation of warrant or grounds runs into a serious problem of trust.

Sometimes truth is equated by semantic non-realists, not with actual recognition of a correspondence between a belief and the world, but with justification in the long run or verification in principle. (I take the phases "in the long run" and "in principle" to be identical). Here the claim is that truth is what will result if we have infinite time and perfect cognitive conditions for our disposal, but such a claim is not plausible. Many descriptive sentences are not empirically justifiable under less ideal but actually realizable conditions, even though we still assume that they are concerned with facts of the matter. Statements like "The law of radioactive decay holds everywhere and all times," "The present universe will last 60 billion years," "No species more cognitive advanced than *Homo sapiens* will ever develop in the future," "69,257 people were killed during the eruption of Vesuvius in 79 AD," or "Until now no two people have been observed to have identical fingerprints" cannot be verified, even if we had enough time and resources. It is reasonable to agree that these claims assert particular facts of the matter and that therefore they may have a definite truth value. Of course, it is quite possible that some theoretical models might predict whether these sorts of claims are true, but nobody would consider such theoretical predictions as a proof that reality has to be as they assert. In order to ascribe a certain truth value to such sentences, we must have observational access to the facts that make them true or false. But in the examples in question, we do not have, and apparently cannot get, that empirical information that would allow us to determine the relevant truth-makers. Nevertheless, to argue that since we are unable to establish the facts of the matter these sentences are neither true nor false seems to conflate the ontological property of being a truth-maker with the epistemic task of proving that it is a truth-maker.

However, as we have seen, the existence of a truth-making fact is only one necessary condition for an assertion to be true. Another necessary condition is that some external fact stands in a certain relation to what is said to be true because of that relation. So perhaps the proponent for an account of truth in terms of verification in the long run could argue that it is not a question of existence of external facts but a question concerning the nature of the truth relation between beliefs and the world.

This truth-endowing relationship exists only if we can, in principle, find grounds for its assertion, i.e., it does not exist only because of the existence of its relata. Such a way of escape is not convincing, for how can we analyze the truth relation in terms of the epistemic notion of justification and at the same time admit that such a justification would require cognitive powers nobody has? Nobody lives infinitely long, not even the human species as such. Even taken collectively as the whole human race, we are definitely very limited cognitive beings in our possible empirical range and scope.

The original idea behind the non-realist move was to make truth accessible to human beings and not to see it as something that would always escape our cognitive limits. The non-realist saw observational phenomena as revealing facts that could be judged by our senses as that which imparts truth to our beliefs and sentences. Unobservable phenomena cannot be objects for our empirical judgment and therefore cannot act as truth-making facts. However, this effort for making truth accessible to humans is all given up again when truth is equated with what can only be reached under such idealized cognitive requirements. In practice the truth relation becomes epistemically unfounded. The realist and the non-realist notions of truth now end with the same result: truth is set beyond the range of what can ever be empirically known. Seeing truth as identical with justification is in every way a mistake: not only does such a project fail our intuitions, but in the end it also betrays its origin by making truth unattainable.

To summarize I have argued that truth cannot be an epistemic notion in the sense that what counts as truth must be defined in terms of an ideal cognitive warrant. Produce whatever descriptive or declarative sentence token you like, if it is true or false, then it is so, not because we possess the ability to know that a fact matches or does not match the sentence, but because either there is a fact or there is not a fact, which makes it true or false, irrespective of our actual knowledge or even possible knowledge. Not only does such an intuition reflect our common-sense realism, but also such a claim recognizes that of course epistemic factors play a role in determining what we can say is true or false. Our cognitive abilities may still have a role to play in contributing to the notion of truth in ways other than just constituting grounds for ascribing truth to individual sentences, as these grounds require both perceptual recognition and various confirmation relations. They may take part in determining what will be the correct sentence type to state in a particular situation in order for that particular sentence to be true.

Thus, the common-sense realist has a good case when it comes to explaining the truth of every empirical statement (belief) as something that is made so by an external fact without any help from our cognition. Such a fact makes a particular sentence true or false because the statement refers to a particular fact. But if the advocate of truth as correspondence understands facts as something given uninterpreted to the receptive mind, he might have overestimated the scope of his own arguments. Only if the metaphysical realist can prove that facts are objectively given as truth-makers and that the truth relation is based on an objectively supplied relation of reference can he claim that our notion of truth is completely mind-independent without involving any epistemic element. In contrast, facts as truth-makers may not be something that can be "read off" of reality automatically and then represented to the mind as mental representational duplicates; their nature may not be as simple as the perceptual identification seems to indicate. Situations, occurrences, and phenomena may not be objectively delimited or demarcated from one another in the world. We do not perceive merely a cat on the mat, one could say. What we perceive is incredibly complex; when we talk of perceiving a "fact" the particular individuals and relations that comprise that fact must be selected from the continuous input of our perceptions. Part of the process by which we identify individuals and relations is determined by natural selection, but part is also arbitrary and depends on our interests and cultural background. Thus, delineating a fact as a truth-maker is somehow dependent on our cognitive framework, epistemic situation, and epistemic interest. We cannot describe or refer to anything that transcends our cognitive limitations and therefore we cannot say anything with empirically determined content about transcendental things. Already Kant was aware of the constitutive character of the conceptual scheme for empirical knowledge, but he also manages to talk a good deal about things-in-themselves—at least as limiting concepts.

In recent times philosophers of science have questioned the objectivity of empirical knowledge as a result of the theory-ladenness of our perception.[12] Apparently, we have to accept that the individuation of facts, at least partly, depends on such conceptual schemes. However, the question remains, of course, how much this admission may ruin the realist's position that truth can be specified in terms of a firm, mind-independent correspondence between a truth-bearer and a truth-making fact. For if such a fact is not something that is (entirely) naturally given, but arises partially from linguistic conventions, how can the metaphysical realist continue to insist, without serious qualifications, on the alleged existence of objective facts? Moreover, if the truth relation is not naturally given as an *internal* relation, how can she maintain this essentially non-epistemic notion of truth?

5.4 A Naturalized Notion of Truth

In order to understand the concept of truth and to explain its epistemic roots in perception, we have to look elsewhere for answers. I shall argue for a natural three-stage model of truth according to which both 'correspondence' and 'coherence' describe different cognitive steps in our entire belief formation. The foundation of truth emerges from causally induced beliefs by our environment long before higher biological organisms developed a reflective consciousness. The formation of beliefs comes with the ability to generate concepts, and this ability exists as an embodied, unreflective cognitive capacity of perception made possible in virtue of natural selection.[13] Nature has by adaptation given us a causal underpinning of correspondence. Those tacit beliefs an organism acquires through its sensory channels correspond causally to some state of affairs outside the organism itself. "Snake, dangerous, get the hell out of here." On this cognitive level of understanding belief formation is part of our causal processing of sensory information and need not be accessible to reflective consciousness in order to take place.

[12] I do not intend to claim that Kant and contemporary constructivists are of a common mind, The whole point of Kant's program was that he claimed to deduce—to prove transcendentally—the constitutive elements as universal and necessary (which secured "objective validity"), whereas contemporary constructivists want to make these constitutive elements relative and arbitrary.

[13] See Faye (2014), Chap. 2.

The next step comes with language and reflective consciousness. Eventually human beings inductively realized by reflection that perceptions correspond to states of affairs that can be distinguished from perceptions. So when human beings developed the capacity of expressing their beliefs in terms of languages they increased their ability to generate new linguistically fostered and transmitted concepts that were not directly connected to the processing of information many, many times. As a consequence, a perceptually accessible state of affairs could be characterized in different ways depending on how the state of affairs was conceptually described in relation to the entire linguistic system. But the cognitive situation is still such that the resulting beliefs are concerned with perceptual states of affairs and intentionally assumed to be true whenever they causally correspond to them.

Finally, the reflective mind generated further concepts that are not directly tied to immediate perception. By abstraction, imagination, and various types of inference, human beings have been able to create conceptual constructions, and to express beliefs in terms of them, that go far beyond sensory information. From an evolutionary naturalist point of view these beliefs are no longer factually true in virtue of a causally established correspondence to some perceptual state of affairs, but they are formally true in virtue of their inferential and explanatory coherence with a network of beliefs. Thus, this three-stage model of truth operates with both a factual notion of truth as correspondence and a formal notion of truth as coherence. So the naturalist would conclude that the notion of truth-makers based on a naturally grounded notion of correspondence has a very different cognitive function than the one based on a notion of coherence grounded in reflection.

The evolutionary naturalist wants to explain the semantic content of reflective thinking in terms of coherence and cognitive representations. The semantic content is not something extra besides cognitive representations and a coherent connection among the representans, i.e., what is supposed to do the representing. A word, a sentence, or another symbol has meaning because it is intended to represent something, and its truth conditions consist in our capacity to establish whether it actually represents what it intends to represent. The semantic antirealist urges that the semantic realist's contention that undecidable sentences do have a truth value is based on a mistaken model of semantic understanding that was originally generalized from decidable sentences to all kinds of assertions. Our understanding of a sentence's meaning has to show up in a linguistic practice as an ability to know whether or not the truth conditions for the sentence are or are not fulfilled. Thus, we should distinguish between those assertions where we can entertain this cognitive ability and those where we cannot. It is this distinction that motivates the evolutionary naturalist to say that whatever a fact is, it cannot act as a "truth-maker" unless it can be recognized to satisfy some truth conditions in the way that semantic antirealism thinks of it in terms of verification/observation/empirical accessibility, etc.

As said earlier, a metaphysical antirealist, who is also an evolutionary epistemologist, is a realist concerning our common-sense world but an antirealist with respect to the meaning of undecidable statements. Moreover, an evolutionary epistemologist, who is a semantic antirealist, may be an agnostic with respect to claims about concrete entities, but a non-cognitivist with respect to claims about abstract entities. The fundamental divergence between the metaphysical antirealist and the metaphysical realist can be formulated as a difference in attitude toward the question of whether truth is partly an epistemic or completely a non-epistemic notion. The metaphysical realist argues that any meaningful assertion has a definite truth value regardless of our lack of cognitive powers to determine which value it is. The metaphysical antirealist denies this, maintaining instead that a particular claim can be true if, and only if, it has some time been empirically or observationally possible to determine the truth relation. Indeed, the metaphysical antirealist argues without difficulty that a necessary condition of being truth conditions is that they are epistemic in nature. This implies that what makes a belief true is the world as it can be observed to be. The metaphysical realist as well as the metaphysical antirealist may accept that the truth-maker exists independently of our cognitive power, but in contrast to the realist, the antirealist takes the truth relation to depend on our cognitive faculties. At the same time this difference between the realist and the antirealist reflects a fundamental disagreement about the features on which linguistic competence rests. Whenever a person grasps an assertion she seems, according to the realist, to grasp the conditions under which a sentence is true or false, but not necessarily to know whether those conditions actually obtain; according to the antirealist we understand the assertion only when we know the conditions under which it can be verified or confirmed.

Characteristic of semantic antirealism is the view that we cannot distinguish between the truth conditions of an assertion and our empirical ability to observe that which would make this assertion true. This view may be associated with a notion of truth as coherence. But semantic antirealism can very well be combined with a concept of truth according to which truth is a correspondence between a statement and the reality it deals with. For the realist agreement between a sentence type and reality exists even in cases where it is impossible to observe or otherwise empirically substantiate what the sentence expresses. The antirealist's claim, in contrast, involves a denial of this: the concept of truth cannot be extended to something that goes beyond our cognitive possibilities of knowing and therefore cannot be applied to undecidable assertions. The difference on this point between the metaphysical realist and the metaphysical antirealist stems from the idea that the realist sees correspondence-with-the-fact as an objective mind-independent relation between the world and a certain statement (and for that reason the truth relation will exist in spite of the fact that we possess no possibility of demonstrating it), whereas the metaphysical antirealist, in the form under consideration, considers correspondence-with-the-fact as a constructed relation trading on the notion of coherence (for which reason truth will be absent if it is impossible for anyone to establish this truth relation). One of the questions then is how such a naturalistic concept of truth can be specified. My contention is that such an explication would disprove that the view of warranted assertibility of statements has to be altogether central for a metaphysical antirealist's understanding of truth. It amounts to a rejection of the view that verifiability in principle has to play a role in an appropriate account of this notion. Nevertheless, such an antirealist truth concept will at the same time explain why truth and warranted assertibility are co-referential but not synonymous.

Earlier I pointed out that the concept of truth is tied to the concept of reference. A descriptive sentence is true or false if, and only if, various parts of the sentence refer successfully to reality. But the issue concerning how meaning and reference are related is a complex and difficult question with no simple and straightforward answer.

On the one hand, it seems clear that if we imagine a physical system of signals, which do not contain any language sign yet, i.e., any sound or mark that purposely stands for something else, then the whole system will have no intended meaning. What is required of a physical sound or mark to become a meaningful sign is that it is intended to reach out for something different from itself. A system without an envisioned reference could not be dealing with anything and would therefore be senseless. A series of objects like 34WcedfV3\ has no meaning unless it is translatable into a known language, or artificially made up, which stipulates certain referring signs, or is causally associated with situations where it denotes specific things. No system of physical marks, like letters, or physical signals, like sounds, can represent something by its own nature. Representing can only happen when we interpret the system to be intentionally pointing beyond the system itself. A physical system can act as a medium for meaning only when we attribute to it certain features, which make the elements of the system refer to mental states or physical elements.

On the other hand, it also seems evident that meaning determines reference. Assume somebody gives a definition of the word "gongo" in the following terms: "A gongo is a small, human-like creature that lives in old caves in the Alps." Such a definite description yields the observational criteria under which the name "gongo" has a referent because the various terms in the sentence provide it with a sense. It is only because we understand the meaning of the words in the sentence that we could take a walk in the Alps and be able to identify gongos if we ever come across them. It would be impossible to connect the word "gongo" with such a creature if we did not understand the content of the sentence. The sense of the sentence makes sure that "gongo" may have a reference, or will have a reference if a gongo exists. A whole language can be used to give sense to new words in terms of definite descriptions, because we already understand the meaning of the language and thus make certain that we can connect these descriptions with the world.

This conflict puts us into a dilemma, for meaning as truth conditions seems to presuppose reference, and reference seems to presuppose meaning as truth conditions. It is the same dilemma we find buried in the discussion about truth as correspondence or coherence. The first scenario fits well into Putnam's remarks about his and Kripke's theory of reference: "The idea that the extensions of our terms are fixed by collective practices and not by concepts in our individual heads is a sharp departure from the way meaning has been viewed ever since the seventeenth century."[14] The second scenario fits equally well with Frege's argument that understanding the sense of a term as it can be explicated by a definite description is what gives us the reference. However, I do not consider the two alternatives as opposed.

From an evolutionary perspective one may argue that many experientially basic concepts were in some distant forefather's head because of all higher organisms' cognitive skills of identifying, comparing and abstracting what they were seeing long before their descendants developed any linguistic practice. It seems reasonable to assume that the first linguistic practice arose assisted by these experiential concepts. *Homo sapiens* simply learned to associate various sounds with kinds of visible objects in their environment. Parts of the cognitive mechanism behind this learning process were that some of the sounds pointed to objects separated from the sound-producing organism. Eventually, as the primordial linguistic system evolved, the reference of singular sounds was supplemented with a syntactical and semantic structure in which coherence among the referential carrying elements gave its users the capacity to understand how the system functioned. This rise of the linguistic system seems to change reference from being a natural result of selection to being mainly a conventional result of intention. But our possibility of establishing the externality of the referent is still confined to the original empirical basis of information.

So I would say what are real are the truth-makers, we are able to observe, and the relationships they have to each other. But how we individuate—in a particular language—the set of truth-making facts that we say characterizes this world will depend on how we 'individuate' the distinct "facts" that comprise this network of relationships. It is quite reasonable to suppose that different languages will individuate truth-making

[14] Putnam (1981), p. 75.

facts differently and so will come up with a description of the world as a different list of facts than another language (non-intertranslatable). Indeed even in one language different interests will lead the fact-lister (i.e., the world describer, a.k.a. scientist) to individuate the truth-making facts differently and so produce a different description of "the world"—many descriptions, one reality; *e pluribus unum*.

Thus, I claim that it is our cognitive power to recognize the referents of the various terms of a sentence that is what determines whether a statement is true or not. It is human beings who establish whether or not the intended reference between a sentence type and the experienced reality is as imagined. Natural reference is not something that exists objectively in nature; it is something that has evolved through natural selection by our cognitive adaptation to an external world. Any later determination of reference in terms of meaning, as Frege suggested, can only be the intended reference, i.e., a reference made up in virtue of a person's access to the entire linguistic system. Determining that the intended reference actually refers to something external—and therefore makes it possible for an expression to be true or not—depends both on our empirical capacity of recognizing the presence of a referent for the term and on our knowledge of the truth conditions of that expression. After this first moment the empirical criteria that allowed us to establish the referent are also those that make the continuous reference intelligible. Moreover, it is these truth conditions that have to be satisfied in order for us to be capable of ascribing a truth value to a sentence of which the term is a part.

As competent users of a certain language we have learned in which empirical situations particular names and predicates can be associated with a referent. By becoming language users we learn which empirical states of affairs have to be fulfilled for using a sentence in order for those states of affairs to be counted as a referent for the sentence in question. And it is only through observation and manipulation of the world around us that we are able to decide whether or not the truth conditions under which the sentence is supposed to have its reference determined are actually fulfilled. Consequently, it is difficult to see how the truth conditions that fix the reference of a given term could possibly be conditions that are not observationally or at least empirically accessible. Here it is important to draw a distinction between the mode of reference and the referent.

What is not at stake is the reality of the referent itself, since it is assumed to exist independently of our actual observation. But what interests us here is the mind dependence of the mode of reference, because the referential relation is constructed in virtue of our natural capacity to recognize the referent as falling under a given descriptive term. There is nothing in the world that makes a particular object to be the natural referent of a given term. On the contrary, we have to realize that the object functions as a referent in virtue of a relation between this word and this type of object that has been fixed according to certain cognitive criteria that are learned as part of learning to use the language. The only way we can find out whether a particular object is the referent of a certain term is to see whether the object stands up to some cognitive procedures that associates this term with this type of object.

Suppose I see a bird outside my window. As an ornithologist I may immediately see that it is a female of the species northern cardinal. But suppose someone asked me how I can be sure that it is a northern cardinal. To respond I must show how it fits those empirical criteria that are associated with the way the reference of the term has become fixed and that eventually convince her that I am correct in referring to the bird as a northern cardinal. I may show her a book that describes the birds of North America in which she can see what a cardinal looks like and that this bird outside my window cannot be mistaken for another bird. The empirical criteria stated in the book are those that determine that it is correct for me to use the name "northern cardinal." Yet, before I have convinced the person that it is correct to call the bird a "northern cardinal," she has to understand the implications of these criteria for the description of the bird sitting outside my window and be able to confirm that the description is satisfied by the bird. If the criteria, that establish the reference of "northern cardinal" were not empirically accessible, she could never actually recognize that the term has a referent.

Thus, it is not until we have discovered the empirical conditions making an intended reference actual that the term can be correctly said to refer to an external reality. Just take the name of the "phoenix." The mythological story tells us that this bird is of great beauty, the only one of its kind, supposed to live 500 years in the Arabian Desert and to rise from its ashes in the freshness of youth and live another cycle of years.

This description provides us with the empirical conditions under which "phoenix" may possibly have a referent. Nevertheless, no observable being in the world meets these requirements. Thus, it is impossible for us to point to a bearer of the name. For the same reason the name has no non-intended reference, and therefore non-analytical sentences about this creature have no truth value. Indeed, a phoenix is presumed here to be observable, but since it is a magical bird, suppose we throw into the definition that it is also invisible to human eyes (like the Greek gods), or it is simply very clever at hiding itself. Then the question arises: Does the phoenix figure into any successful scientific explanations, and does it produce any observable effects? So even if we cannot see a phoenix, we might still conclude that, like protons, it is real but invisible.

Thus, if the truth conditions fixing the reference are observationally judged to be successfully satisfied in a situation, when a certain sentence is put forward, we can reach the conclusion that the description is true; if not, we are in a position to claim that it is false. Let us illustrate this with a couple of examples. Both a sentence like "The northern cardinal lives in America" and a sentence like "The proton has an electric charge" meet the possibility conditions for being true. For anything to be a cardinal it must match certain experiential criteria making it correct to call it a cardinal. It has to be a bird having such and such an appearance, such and such a song, and such and such a behavior and habitat. Those are all perceptual criteria allowing the name "cardinal" to have a meaning. At the same time we can determine through observation whether or not there is something in the world that fits these conditions. By these conditions we can tell whether or not the name refers to something in the world. The same applies to the name "America." So if bearers of the two names can be shown to exist in a relation as expressed by the above sentence, then tokens of that sentence will be true, otherwise not.

Turning to the other sentence, "The proton has an electric charge," I argue similarly that such an expression about invisible entities can be endowed with a truth value only if it can be proved that there is a bearer of the names cited in the sentence. Thus, any token of that sentence is true under the assumption that the names "proton" and "electrically charged object" can be shown to refer to the same entities. In fact, through a well-established experimental practice science has stipulated the observational conditions for something to be regarded as the bearer of these names. Certain definite experimental procedures have been stipulated to lay down the experiential conditions under which something can be correctly called a proton or an electrically charged object. One might say that something is a proton or an electrically charged object if, and only if, it fulfills those observationally determined truth conditions for one of these names to have a reference.

As a consequence, statements of the above sentence are true if, and only if, positive experimental results appear when we perform the relevant procedures for identifying the referents of the names. However, it is also quite obvious that an evolutionary epistemologist cannot ascribe truth value to sentences about invisible objects and properties unless there are certain experimental procedures guaranteeing a reference of the descriptive terms. It also implies that the evolutionary epistemologist can admit modal, universal, and counterfactual statements have a truth value only to the extent to which the descriptive terms contained in these statements can be related to states of affairs to which we have observational access.

Thus, my claim is that the truth relation is an *external* relation rather than an *internal* one. An assertion is automatically true only if it can be decided that its truth conditions are satisfied. The existence of both the truth-bearer and the truth-maker has to be individually acknowledged before the truth relation is in place. An individual statement is true if, and only if, (1) it corresponds to the stated fact, and (2) this correspondence once has been established by a *descriptive correlation* between this type of statement and this type of fact because we as cognitive beings have been able to identify empirically (directly or indirectly) the sentence type and the fact type in isolation from each other. This constructivist aspect of truth underlying metaphysical antirealism excludes the possibility that at the same time it is meaningful to talk about the world as it is independently of our observational possibility of establishing the truth relation.

In my view there are several reasons for defending such a notion of truth. First it seems to be consistent with the claims of evolutionary epistemology. Second, it embraces the full consequences of the general challenge that Dummett has raised against semantic realism.[15] Third,

[15] This is the great theme of Dummet's investigation into philosophy of language. See, for instance, Dummett (1973), (1976), and (1978), where he develops his arguments for semantic antirealism.

such a theory is able to explain how we are able to learn our first language. Some words can be learned by definition or translation, but this requires that we already have learned a language. The problem for the semantic realist, who holds the existence of verification-transcendent truth conditions, is that he allows that an undecidable sentence can be true even though a competent user of that language cannot know whether the truth conditions are satisfied or not. A semantic antirealist does not fall victim to this problem because her truth conditions are not verification-transcendent. This implies that a language can be learned by induction based on empirical information.

Therefore, I urge that the debate between scientific realism and empiricism changes character. The old ontological question of whether invisible entities exist or not should no longer be central to the dispute. In opposition to the empiricists, I think it is possible to be an ontological realist concerning invisible entities but a semantic antirealist concerning their truth conditions. As long as the realist and antirealist think of their differences in terms of ontology concerning the reality or non-reality of invisible but observable entities, modern science and technology give realism an easy victory. This may also explain why some avowed realists, like Michael Devitt and Ernan McMullin, claim that realism is not connected with truth, but that ontology has to be separated from semantic.[16] The current dispute, however, should be a debate about the conditions under which a description of any entity can be said to be true or false. With respect to a reformed debate, the realist's triumph turned out to be a Pyrrhic victory.

5.5 Semantics and Ontology

Until now I have argued that metaphysical realism consists of an ontological as well as a semantic component, which are somehow connected, and these two components are also part of metaphysical antirealism. It has been argued, especially by Michael Dummett, that an equivalence exists relating semantic and ontological realism on the one hand and semantic

[16] See Devitt ([1984]1991) and McMullin (1984).

and ontological antirealism on the other.[17] Semantic realism argues that our understanding of the meaning of a declarative sentence consists in knowing under which conditions it is true or false. The realist insists on the meaningfulness of ascribing a definite truth value to declarative sentences regardless of our inability to determine whether or not its truth conditions are satisfied. Such sentences are true or false even if we do not possess any procedure by which a truth ascription can be determined. Then, as Dummett's argument goes, the claim that such statements have a determinate truth value is the same as saying that the reality described by these sentences exists independently of our ability to discover anything about it. So the semantic realist must also be an ontological realist.

In contrast to the realist view, Dummett argued that the antirealist view involves two opposite components.[18] Antirealism denies that it makes sense to ascribe a truth value to sentences if we cannot decide that its truth conditions are satisfied. The antirealist makes a distinction between (effectively) decidable and undecidable sentences: decidable sentences are those declarative sentences for which we are able to know whether or not their truth conditions are satisfied, whereas undecidable sentences such as modal, universal, conditional and counterfactual sentences are those declarative sentences where we have no (empirical) means to decide that their truth conditions are satisfied or not satisfied. However, there is a snag with Dummett's characterization of what is decidable: the truth conditions have to be effectively shown to be satisfied in terms of actual verification. This is much too strong. Dummett would say, for instance, that sentences regarding possible historical facts such as "Caesar slept on his right ear on the night of his first birthday" have no definite truth value because it is reasonable to assume that today no evidence exits by which we can settle that claim. Instead of Dummett's actually verifiable criterion, I would argue that such sentences are in principle verifiable since had we been present at Caesar's first birthday, we would have been able to experience what was the case. The situation is equivalent for distant

[17] Dummett (1973), Chap. 5 and Chap. 6, as well as Dummett (1976), Sect. VI.

[18] Cf. Dummett (1973), 464–470; Dummett (1976), pp. 81–82; Dummett (1978), pp. xxvii–xxxiv.

events even if they do not leave any traces that can count as evidence. Such a suggestion is more in accordance with our realist instinct.

From an ontological point of view, many metaphysical realists will see the satisfaction of the truth conditions in terms of truth-makers in the sense that they assume what has been called truth-maker maximalism.[19] For all propositions <*p*>, <*p*> is true if and only if *p* makes it true. The motivation behind truth-maker maximalism is that if some propositions are true in virtue of truth-makers, there have to be strong reasons to refuse to apply truth-makers to all true propositions. In his rejection of this maximalist view, Dummett's semantic antirealist argues that the only propositions that are true are those that can be actually verified, and about those we may say that they are true in virtue of discoverable truth-makers. However, an evolutionary naturalist sees the matter in a different light. Take the above counterfactual statement, "Had we been present at Caesar's first birthday, we would have been able to experience what was the case." As a generous metaphysical antirealist, the evolutionary naturalist would claim that such a statement is indeed true, albeit no truth-maker makes it true. This is so because we inductively infer the truth of some counterfactual statements from true propositions about which we have learned by experience that in the right circumstances we can discover their truth-makers to see whether or not they make the corresponding statement true or not. Hence, some statements may be true or false by correspondence, some are true or false by coherence, and some are neither true nor false. Which ones fall in which class all depend on our cognitive resources.

It is essential for the semantic antirealist that linguistic meaning is manifested in the linguistic practice by which the speaker demonstrates his or her implicit knowledge of the use of a given expression.[20] This implies that we must be able to show that the speaker understands

---

[19] See, for instance, Armstrong (2004), Rodriguez-Preyera (2005) and Cameron (2008). They all defend truth-maker maximalism. Armstrong says: "My hope is that philosophers of realist inclinations will be immediately attracted to the idea that a truth, any truth, should depend for its truth on something 'outside' it, in virtue of which it is true" (p. 7). Similarly, Rodriguez-Preyera claims that "the root of the idea of truthmakers is the very plausible and compelling idea that the truth of a proposition is a function of, or is determined by, reality" (p. 20).

[20] Cf. Dummett (1976), p. 80, and Dummett (1973), pp. 460–462.

the meaning of a sentence by showing that we possess a (effective) method for deciding whenever its truth conditions are fulfilled.[21] As Dummett pointed out, we cannot meaningfully ascribe implicit knowledge of the meaning of a given sentence to a person unless it is possible to specify what the manifestation of this meaning consists in. The antirealist would say that the realist wrongly believes that we have managed to supply our sentences with such a semantic content that every statement is either true or false. This assumption misleads the realist in the case of undecidable sentences to believe that these phrases have a content determined in relation to circumstances beyond our cognitive abilities to establish (in principle). Thus, if this objection is correct, it is meaningful to ascribe a determinate truth value to a declarative sentence only if it could in principle be verified, or otherwise confirmed, that such circumstances obtain. According to the antirealist view, the truth of a statement is dependent of our ability to recognize the fact of the matter, but this in turn implies that the truth-making reality becomes dependent on our cognitive powers, because with respect to undecidable sentences we cannot determine whether or not the situation fulfilling their truth conditions occur. Therefore, it is we who decide what counts as real and what not and therefore which sentences possibly can be either true or false. If this argument is valid, it seems that there is little room for a clear distinction between ontological and semantic antirealism, owing to the fact that semantic antirealism entails ontological antirealism, and vice versa, just as ontological realism entails semantic realism, and vice versa.

As a proponent of evolutionary epistemology I want to maintain that there exists an external world that endows beliefs and statements with a truth value, but the meaning of a sentence is not established by its truth conditions but by its verification conditions. The semantic antirealist does not have to be an ontological antirealist, because when talking about decidable sentences it is possible to get to know what makes them true or false by pointing to external states of affairs as what satisfies their truth conditions. Unless such sentences are about our mental states, their semantic content is about external things to which we have observational access and whose discovery we can use to reach a definite conclusion that

[21] Cf. Dummett (1978), pp. 216–217.

a situation making the sentence either true or false occurs. So the semantic antirealist will refuse that it makes sense to claim that undecidable statements have truth-makers.

As Dummett saw the commitments of realism and antirealism, they have no bearing on the realist-idealist distinction, and I see no way in which they can be revised, so an evolutionary ontology that is committed to ontological realism and to semantic antirealism is excluded. The suggested strict equivalences between ontology and semantics simply do not hold. In order to see why, we shall first look at the implications with respect to metaphysical realism. The argument for a biconditional between semantic and ontological realism presupposes that the principle of bivalence, i.e., the claim that every well-formed sentence is either determinately true or determinately false, is a defining part of metaphysical realism. But though it has been, and perhaps still is, part of the standard definition of many philosophers, of whom Dummett is the most significant proponent, there are no good reasons from a metaphysical perspective to maintain that the validity of the principle of bivalence is necessary for being a realist. Likewise non-realism cannot be associated with the universal denial of bivalence.

A better way of defining semantic realism would be to say that contrary to semantic antirealism such a position just requires that truth may transcend any means of verification, which implies that a belief or a sentence can be true or false even if it cannot in principle be proven to be so. This definition of realism in terms of verification-transcendent truth conditions (and not bivalence) allows the semantic realist to hold that some well-formed sentences within a certain domain are neither true nor false, either because the objects or events to which these sentences appear to refer are unreal or are real, but have vague or indeterminate properties. For instance, a metaphysical realist can think of the future in five different ways: as real and determinate, as real but indeterminate, as only partly indeterminate, as partly unreal or as completely unreal. Both the second and the third assumption imply the objectivity of becoming, but, nonetheless, there is an important ontological difference between them. The difference I want to point out is one that appears in the distinction between the reference of names and the satisfaction of predication. A sentence about, say, a future event, like "*X* is *F*," can fail to be true (or false) for two reasons: (1) "*X*" does not refer to any event in the future. This is the case in which I take the future to be unreal; (2) the event to which "*X*" refers exists, but the ascription of the property *F* to *X* is not quite true; neither is it quite false. Instead, such an attribution is possibly true and, therefore, also possibly false. This is the case in which I take the future to be indeterminate.

We can then give a characterization of two metaphysical realist positions with respect to the future, both of which accommodate the semantic principles of verification-transcendence:

The first position claims that future sentences have no truth value whatsoever; not only does it reject the universal validity of the principle of bivalence, but also it actually insists on the validity of the negation of the principle claiming that there exists a declarative sentence about the future, *p*, such that *p* is neither true nor false, i.e., (1) $\vDash \exists p \sim (\mathrm{T}\, p \vee \mathrm{F}\, p)$. The underlying idea is that some, or all, future sentences are neither true nor false because they completely lack any truth value. Such sentences do not even have a truth value with a degree of probability. This is due to the fact that the future is (partly) unreal; no facts exist to which these sentences may refer to endow them with a truth value. Moreover, in case the metaphysical realist also thinks of the past as unreal, as the proponent of presentism does, he must argue that objective facts themselves are tensed. For the proponent of such a view truth cannot be stable, but will change as the facts change accordingly from being unreal to being real, and then from being real to being unreal again. A gained truth value, gained only when a possible fact becomes actual, will not last forever. The very same token, "It is now raining," which became true yesterday, may be false today. This assumption is in conflict with the basic principle that a proposition is always true, if it is true at all.

The second position, however, merely refuses to accept the universal validity of bivalence, the principle that every declarative sentence *p* is determinately true or determinately false, i.e., (2) $\nvDash \forall p\, (\mathrm{T}\, p \vee \mathrm{F}\, p)$. Rather some sentences about the future have a truth value with a certain probability. Here the underlying idea is that the principle of bivalence is not universally valid because some future sentences are only true or false with a certain probability, owing to the fact that the future is (partly) indeterminate.

It is worth noting an ambiguity in different possible understandings of probabilities. If we just state "$P(Ax) = 1/n$," we can either construe such a statement in terms of bivalence or interpret it as an expression involving probabilistic truth. The common reading would be that "'$P(Ax) = 1/n$' is true if, and only if, $P(Ax) = 1/n$," and "'$P(Ax) = 1/n$' is false if, and only if, $P(Ax) \neq 1/n$." In this case the realist would think of probabilities as dealing with determinate facts associated with a class of objects. For instance, the probability may designate the relative frequency within a class of objects, of which x is a member, that has the property A. A not so common reading, but possible anyway, could be to understand the probability statements as "'Ax' is true with the probability $1/n$ if, and only if, $P(Ax) = 1/n$," and "'Ax' is false with the probability $1\text{-}1/n$ if, and only if $P(Ax) = 1/n$." In that case the realist would regard probabilities as something that concern indeterminate facts of objective chances and propensities associated with individual objects.[22]

Thus, in my opinion, the metaphysical realist is not forced to hold the validity of bivalence, and whether she actually accepts or rejects bivalence seems to depend on the nature of the domain of reality with which his investigation is concerned. Assuming that the semantic realist takes the existence of atomic objects as something that is essential to the truth value of sentences in quantum mechanics, she could argue with respect to atomic objects that some of their properties are always vague or indeterminate, while others are precise or determinate.[23] But what does that exactly mean? The canonical formalism contains n pairs of non-commuting variables representing n pairs of incompatible properties. A complete description of a quantum system, however, requires at most that only one of a pair of n commuting variables can be assigned a sharp value

[22] The use of probabilities in this discussion is meant as an illustration. The pretty much inescapable fact is that we use probability statements in a variety of ways, and it seems to me inadvisable to try to shoehorn them all into a single interpretation of probability statements. There are clearly many times when we say the probability of X in circumstances C is $1/n$ as a simple induction from a record of frequencies, we simply add up all the incidents where C obtained and where X did or did not occur, and calculate that percentage of the total. Here probability is purely empirical, and all evidence on which a judgment is made is empirical. If I am absolutely in ignorance about the truth condition of probability statement P, I may say the probability of P is true is 50 % purely as a point of logic. When the weatherman says the probability of rain in Zealand tomorrow is 50 %, it is neither a logical truism nor an induction from past frequencies, but a matter of deduction from the physics of the atmosphere as best (but imperfectly) understood by meteorological science. It is also relevant to point out that there is no calculus of inductive probability.

[23] See Petersen (1985), Rohrlich (1986), and Krips (1987).

simultaneously. The system is brought in the situation where it has the power to confer n sharp values to the variables whenever the respective attributes are measured. The other n variables do not have sharp values, because the properties that these observables stand for are indeterminate until they are measured. Thus, the vagueness of a quantitative attribute means that a system only indeterminately possesses this property, since the property may have very possible value and therefore only has a definite value with certain likelihood. This is described as if there are objective propensities for the system to manifest a certain value. The words "vague" and "indeterminate" would refer to the fact that the atomic object only holds a certain quantitative attribute with a certain objective probability. If this is correct, it amounts to saying that each sentence that ascribes a vague attribute to an atomic object does not have any determinate truth value, because a mind-independent or objective propensity exists in the system that may manifest every possible property value.

In his understanding of vagueness the metaphysical realist seems to have two options: he could either say that statements ascribing vague attributes to an object have determinate truth values, because objects having vague properties create the truth conditions for having a determinate truth value, i.e., it is definitely true that the system has vague properties. Or he could argue that statements ascribing sharp attributes to an object may be true with a certain probability, because atomic objects in such cases possess particular sharp property values only with a certain probability. In the latter case sentences about vague objects may still, according to the realist, have semantically well-defined verification-transcendent truth conditions, but the truth function assigning a truth value to the sentences on the basis of this objective tendency is only partial. So, the fuzzy advocates are not saying anything about observed properties, which they know always give sharp values; they are claiming the *unobserved* parameters have only fuzzy values, and by definition the truth conditions of such statements are verification transcendent. In other words, the world may well exist objectively, entirely independent of our cognitive apparatus, while parts of reality have such a nature that they do not endow every declarative sentence predicating sharp values to atomic objects apart from their observation with a certain truth value. In the latter case, however, we realize once again that the semantic realist supports a view that is in contradiction with one of the basic principles of truth according to which something cannot be more or less true. If the truth predicate is ascribed to a sentence with greater or smaller probability, truth cannot be absolute.

So much for the metaphysical realist's possibility of blocking the implications between semantics, in case it is defined in terms of bivalence, and ontology. Like realists, however, antirealists can also hold different semantic positions depending on their view concerning the principle of bivalence. We can say that the semantic antirealist may refuse to accept the universal validity of bivalence, the principle that each and every assertion *p* is true or false, namely in those cases in which we actually have no observational warrants to assert or to deny *p*. In contrast, we can also say that the semantic antirealist may not only reject the validity of the principle, but actually insists on the validity of the negation of the principle by saying that there exist sentences *p* that are neither true nor false, namely in those cases of undecidable assertions that allegedly refer to states of affairs that we cannot in principle possess any perceptual grounds to know. Formally, we can represent the first form of semantic antirealism as $\nvDash \forall p(\mathrm{T}p \vee \mathrm{F}p)$, a position I call semantic agnosticism. In contrast, we can represent the second form of the semantic antirealism as $\vDash \exists p \sim (\mathrm{T}p \vee \mathrm{F}p)$. I shall call this form semantic non-cognitivism.

Thus, in philosophy of language the discussion that was begun by Michael Dummett has in my opinion entirely disregarded the distinction between agnosticism and non-cognitivism. As an example, let me just mention the discussion about the reality of the future. Here, on the one hand, a semantic agnostic can very well accept that the future is possibly real provided, and only provided, we can empirically justify, in some way or another, present statements referring to future events (for instance via backwards causally produced evidence). On the other hand, the semantic non-cognitivist will completely reject the idea that the future is real, since it is impossible to have perceptual evidence to confirm any present statement referring to a future time, and he will therefore rule out any possible suggestion that backward causation is possible.[24]

It is correct that there is an essential disagreement between semantic agnosticism and semantic non-cognitivism in the sense that, given the

[24] See Faye (1989), pp. 86–89.

right circumstances, agnostics may allow themselves to believe in the existence of entities that are excluded from being seen by our normal vision, but non-cognitivists prohibit themselves from such metaphysical extravagances. Since both are versions of antirealism, it is a matter of fact that the assertion of sentences as true or false depends on our ability to ascertain what sentences are about. There is, obviously, an important epistemic consequence between whether one, as an agnostic, overrules the validity of the principle of bivalence for some sentences, or whether the other, as a non-cognitivist, denies the validity of the principle in virtue of accepting the argument for the validity of the negation of the principle with respect to these sentences. Thus, semantic agnosticism gives expression to the view that a certain group of sentences can actually be said to be neither true nor false. The agnostic is agnostic about whether sentences concerning objects, about which we have no actual grounds for ascertain their truth or falsehood, do have in fact any determinate truth values. Since she denies that the principle of bivalence is valid in all cases, she is merely saying that for the moment there is no warrant for asserting or denying *p*. But it may still be possible that *p* at some future time could prove to be bivalent. The semantic non-cognitivist, in contrast, is not an agnostic, for he claims that *p* entirely lacks any truth value. He excludes on epistemic grounds that *p* could become bivalent at any future time. In other words, the semantic non-cognitivist believes that there are domains of well-formed sentences that it is never valid to ascribe a determinate truth value, whereas the semantic agnostic holds that although bivalence may fail to apply to some sentences for the present, it is possible it may become applicable at some future time as human cognitive reach extends into new domains.

Although they differ on the principle of bivalence, the defining characteristic of both semantic agnosticism and semantic non-cognitivism is that meaningful sentences must be decidable by having truth conditions that are empirically accessible, at least in principle. Only by being able to recognize the circumstances under which we can claim that an assertion is true or false can we explain what we have to know in order to understand the content of the assertion. But how does this affect the alleged equivalence between semantic and ontological antirealism? How can this view avoid a contradiction, when, on the one hand, it claims that there is an external world independent of human beings and, on the other hand, argues that what can be said truthfully about the world is dependent on our epistemic abilities? Thus, one might believe that if an objective world exists, then it is the factual state of things that, independently of our cognitive power, determines what is true.

Crispin Wright points out that one traditionally associates objectivity with an assumption that an assertion is true independently of the way we establish that it is true.[25] On the one hand, we distinguish clearly between our beliefs and opinions concerning a particular case and, on the other hand, what is true about this case. Neither our opinions themselves nor the manner in which they are acquired decide what truth is. The notion of objectivity involves at least the idea that a statement, or at best a certain class of statements, can be assigned a truth value previously to and irrespective of any actual investigation. Wright suggests that truth values are investigation independent.

This distinction seems to make room for an evolutionary naturalist's belief in an external world without coming into conflict with his supposition of an epistemic component of truth. However, the term "objectivity" has several meanings. In case one takes "objectivity" to denote a reality that exists independently of any *possible* investigation (by everyone), the evolutionary epistemologist would say that she cannot make much sense of such a notion where a certain class of statements is regarded as true or false regardless of our ability to establish its truth relation. But if one associates "objectivity" with a class of statements that is made true or false by a reality independent of anyone's actual investigation, the evolutionary naturalist has no objection to such a discourse. One just has to maintain that a sentence possesses a truth value if, and only if, the stated fact of the matter can in principle be ascertained by experience. An assertion must, according to the evolutionary naturalist, be empirically decidable in order for it to be possible to say that it has a truth value.[26]

---

[25] Wright (1987), p. 5 and pp. 148–49.

[26] I intentionally use assertion here. A "declarative" sentence is a kind defined in grammar, ending in a period. Not all declarative sentences make assertions, although that is their ordinary purpose; however, there is no single matchup between the grammatical categories—declarative, interrogative, imperative and exclamatory—and the purposes sentences may serve, and of course sentences may serve multiple purposes, and in any case the purpose a sentence may serve cannot be decided

Accordingly, let us turn to the assertion that semantic antirealism implies ontological antirealism. A semantic antirealist traditionally champions the assumption that the truth of a statement is determined by the epistemic grounds for asserting it. The reason is that our grasp of that statement must be tied to its use if some posited feature of its meaning should not remain unknowable to the speaker (and the hearer). Moreover, as Dummett emphasized, we seem to learn a language inductively through the association of the use of certain sentences with the occurrence and non-occurrence of certain empirically recognizable situations, and it is only on the basis of our use of these sentences appropriately in such situations that other speakers of the language can determine that we really understand the meaning of the sentence. However, the antirealist need not insist that the truth conditions of a certain class of sentences must be analyzable in terms of our beliefs in order for particular sentences of that class to be true. Instead she may argue that there are external facts that make decidable assertions true or false, but that the truth relation does not rest entirely on non-epistemic processes. Thus, she could argue that these truth conditions have to be specified in terms of a possible state of affairs that one can have observationally warranted beliefs about, while claiming that this external state of affairs may exist independently of our cognitive apparatus. As a semantic antirealist the evolutionary epistemologist simply denies that sentences about material objects, perceptual or non-perceptual, can be translated into sentences about our opinions. Truth conditions of material object sentences cannot be specified in terms of our epistemic convictions but at most in terms of what is in some way or another accessible to experience. The antirealist may therefore argue that what makes a decidable sentence about a material object true or false exists in the world independently of any actual investigation. But he also holds that this reality endows the sentence a truth value only because what makes it true is acting as a 'truth-maker' in virtue of its relation to the history of cognitive human beings.

out of context involving reference to both speaker and audience. However, by definition, only those sentences that make "assertions," no matter what their grammatical form, have truth value. So I'd recommend avoiding the declarative grammatical category altogether.

Thus, something is a truth-maker if, and only if, in the first place it has been ascertained to satisfy some truth conditions.

The antirealist contention under discussion is that nothing can act as a *truth-making* fact unless it can be empirically or observationally grasped as such. Our sensory and instrumental understanding of the world makes it what it is. The way we pick up the world by our senses or by our instruments determines what truth-making facts are for us: what counts as a truth-making fact, and what not, depends on the categorization we make and on the language we use to state those facts. Where the realist in general believes that the world is once and for all divided into different compartments of natural kinds and categories by nature herself, the antirealist considers it as a conceptually structured reality where the categorical divisions are made by humans. For such an antirealist, of course, this does not mean that the reality in itself is an amorphous and non-structured whole, since, from his point of view, it makes no sense to distinguish between how the world is uncognized by humans and how they conceive it.

Indeed, both the realist and antirealist can agree that the categories employed by any language used to express facts (make truth-valued assertions) are a human construction, because we know that not all human cultures and periods of history employ the same categories. But at least for those scientific realists who are naturalists, over time this categorical scheme evolves so that it more and more closely matches the actual categories of reality, as Plato said, it "cuts nature at the joints." So there are linguistic (or subjective) "categories" and ontological (objective) "categories" (a.k.a. "natural kinds"). But the evolutionary naturalists would not typically argue that there is positive selection in favor of organisms employing cognitive schemes in which the former approximate the latter. Instead, as semantic antirealists, they would maintain that our cognitive schemata are a result of human adaptation to the empirical world. However, the ontological antirealist of the radical social constructivist type has to deny this, because for her there are no ontological (objective) categories; even if she accepts that an independent reality exists, she has to insist that nature does not come self-categorized.

Thus, I shall advocate semantic naturalism different from both semantic realism and semantic antirealism. It maintains the principle of extensionality, which says that the truth of any extensional statement, *A*, is solely determined by the actual situation. But not all extensional statements such as universal statements are empirically decidable. An extensional statement *A* can be decided to have a truth value if, and only if, we are able to observe whether or not its truth conditions are satisfied, i.e., to determine whether or not *A* corresponds with a truth-maker. This excludes some true extensional statements from being true in virtue of a truth-maker. Furthermore, other statements that are either true or false in relation to situations other than the actual situation are intensional, but no truth-maker exists that makes them true or false. These true statements without a truth-maker are claimed to be true by inference. Since undecidable statements do not address the current situation their truth does not come from any correspondence with a truth-maker but from their inferential coherence with all other decidable true statements. Finally, any statement, whether extensional or intensional, that cannot be determined to be true or false by either observational means or by inference from true extensional sentences, are neither true nor false. For instance, negative statements are not true in virtue of truth-makers, but because we deductively infer their truth from positive, decidable statements; modal statements are not true in virtue of truth-makers, but because we inductively infer their truth from positive, decidable statements; the same holds for universal statements. All decidable sentences are indicative, i.e., non-modal, whereas at least some modal statements refer to possible states of affairs that are "contrary to fact."

So the semantic naturalist argues that the conception of truth cannot be explained exhaustively by referring to the correspondence theory. Both the semantic realist and the semantic antirealist may find support in this theory. However, the former is characterized by the fact that she separates truth conditions and the conditions under which the ascription of a truth value can be warranted in such a way that it is always possible for an undecidable sentence to be false though our most informed judgment concludes it is true, and vice versa. In other words, the semantic realist claims that even if an undecidable sentence would be characterized as true after perhaps infinitely many perfect investigations, it might in principle still be false. This is because various parts of an undecidable sentence have a unique intended reference, irrespective of the impossibility of establishing definitively whether this reference relation is successful or not.

For the semantic naturalist, however, it is not entirely possible to separate truth conditions and the ability to substantiate what counts as truth-makers. If it is correct that truth depends on reference and representation, one could argue that the truth conditions assigned to a class of sentences rely on the innate conditions of reference, and these conditions, yielding the truth relation, have to be accessible directly or indirectly to the senses of humans and their instrumental investigation to see what such an interpretation of meaning is all about. The argument is that a certain sentence type can be true only if the various parts of it refer to actual elements of reality and only if the question of whether they refer or not can be settled by observational or operational criteria. In other words, a term cannot be said to have an external reference unless it is possible to fix the referent of the term by certain experiential means. For instance, a term like "God" does not, according to the metaphysical antirealist, refer to a certain external entity because we have no observational methods to establish its truth conditions and thereby its reference. It seems possible, by explicit definitions, to introduce as many terms into a language as one likes, and by doing so, we can formulate meaningful sentences containing these terms. Nonetheless, this does not thereby ensure the externality of their extensions.

6

Abstraction and Reification

What do we impose upon our experience and what can be drawn from it? It is well known that Hume argued that experience cannot be the source of beliefs in causal powers or physical necessities. All that is directly given by sensory experience is the constant succession of similar sensory impressions in time. However, the mind is disposed to see necessities in them as soon as it recognizes a few repetitions of a pattern of events. It is the mind that reads those necessities into nature. Likewise, Hume maintained that our belief in the constancy of objects is due to a similar psychological propensity. Beliefs in the permanency of the everyday objects in our surroundings go well beyond what our senses allow us to experience. Hume granted that the human senses provide us with some idea of object's stability and coherence, but that we assume much more while trusting that objects continue to exist regardless of whether or not anybody is looking. The belief benefits our purposes, Hume said, although it is the human mind that enforces its notion of coherence and stability on the experience of things.

For obvious reasons Hume was not defending a Darwinian epistemology but his was as close as it could be as the immediate progenitor. Also, according to the Darwinian, humans are disposed to impose their cognitive schemata and epigenetic rules upon the information received through

J. Faye, *Experience and Beyond*,
DOI 10.1007/978-3-319-31077-0_6

the senses. As Ruse remarks, "Problems for the Humean are problems for the Darwinian."[1] Not quite, however. Hume was not able to tell us why the mind possessed psychological dispositions to mold and structure our experience by certain categories that were not traceable back to the experience. The Darwinian is in a different position. She could say that the notion of physical necessity and that of the permanency of physical objects exist as innate categories, because it has enhanced the individual's chances for survival and reproduction if it has a propensity to see necessity in succession. Those humans who believed in causal powers were biologically better off than those who saw only accidental relations. It is because such beneficial adaptations have evolved by natural election that we think of causation in terms of necessary connections.

The evolutionary naturalist cannot rest her case here. What happens to the skepticism that kills Humean epistemology? Is the real world just a figment of our imagination that has proven its value in the fight for survival? It is easy to see what has happened to those humans who did not believe in an external reality that common sense teaches us. They are long gone. Eaten by animals (which did evolve with an implicit belief in an external world) or died by accidents. But do we simply assume there is a real world just because believing this has been useful for our predecessors or because in fact there is an external world out there? The short answer is both. It is impossible to image that we as well as our distant predecessors evolved with the cognitive adaptation of believing in an external world if this world is merely a figment of the mind.[2] Assuming that the senses somehow tell us about an external world separated from us, the next problem becomes where we shall draw the distinction between categories of the mind that are there as a result of natural selection and categories that we have consciously created in human history and then imposed on nature herself? This leads us to reconsider some of the key concepts in science that may rest on abstraction and reification.

The main suggestion of evolutionary naturalism is in line with Hume's view that there are no abstract entities, neither particulars nor universals, and that some of the things we take for real, such as numbers, in fact are

---

[1] Ruse ([1986]1998), pp. 185–186.

[2] As well as a good part of the animal kingdom. If we can speak of animals as having beliefs, then surely the belief that there is an external world must have appeared fairly early in animal evolution. Even if they do not have beliefs per se, they surely have instincts that are predicated on the presumption of a real external world.

abstracted from the senses and later “reified” or “objectified” by our mind. Because the realist instinct has survival-enhancing benefits, reification is so attractive and seductive to all of us. We have gained the ability to abstract from our sensory information about concrete particulars in virtue of certain cognitive mechanisms. The explanation is that it has been of great evolutionary benefit for us to grasp our actual experience in ways that allow us to understand it to be similar to other earlier and later experiences and understand what might have led up to this experience and what might follow from that experience. However, the process of abstraction works in at least two ways. One happens automatically or subconsciously because our inborn cognitive mechanics do what they are adapted to do without being informed by any conscious reflection; the other is not carried out automatically, but is a result of a conscious work by our imagination and reflection. In addition to the ability of abstraction, we have also gained an ability to project these abstractions onto our sensory experience. Such a projection is part of the biological value of abstraction. It is from such a natural projection that we gain the habit of reifying those abstractions to make them self-subsisting parts of the real world. This trick originally evolved as parts of the adaptation of our predecessors’ cognitive capacities, but it apparently continued to be the operative mechanism when the evolution of imagination and reflection took place. However, we should not forget that the projection of abstractions is due to our mind and that in the hands of reflection it becomes a misleading guide with respect to metaphysical thinking.

6.1 Common Sense and Externality

The Darwinian scenario pictures the failure of the organism to ‘fit’ its environment as that which causes to perish before it can reproduce. As a scientific assumption, Darwin’s model of selection presupposed that a world exists that is separated from the organism itself. This model explains the organism’s environment as the key causal factor in the evolution of new organisms and new cognitive adaptations of organisms. An organism’s surroundings determine which cognitive schemata and mechanisms (or epigenetic rules) are useful and which are not. So the question about externality boils down to the following issue: do humans have a notion of an external world only because this notion has functional value or do they also have it because there is an external reality? An obvious response would be to say that the notion of an external reality has a functional value for us, because no concept can have a functional influence unless that which falls under it exists. But if this is true, it implies that the same holds for causal powers and physical necessities in the world. They must also exist because of their role in the ability of humans to interact with natural processes. A closer look also reveals that beliefs can be extremely useful without standing for anything in reality. From a Darwinian perspective people have religious feelings because these emotions create social cohesion and help followers to reconcile and find comfort in the midst of a miserable life. Even accidental but false beliefs may sometimes be useful for human actions by preventing people from doing something with fatal implications. So beliefs may be functional also when they are far from true. What matters is that one has faith in them.

A different response would claim that our cognitive schemata and mechanisms have been 'selected' by nature according to whether they give human beings the capacity to experience nature as it really is. Eventually our thought processes have adapted to disclose our physical surrounding as objectively as it can be. Again such a suggestion does not hold up to closer scrutiny. Natural selection and adaptation do not bring us nearer to the ultimate truth. We have the cognitive schemata we have from natural selection since they were genetically available for our predecessors as a way of handling the world. These schemata might have been different, in which case our thought processes would have been different. Nature does not wait for a perfect match between the beliefs, which the genetically inherited schemata lead us to assume, and the way the world really is. Selection works on the material given at any time, and an organism has to make the best adaptation out of it. So from an evolutionary perspective the way we experience the world is contingent; it is no more 'correct' or 'necessary' than any other adapted form of cognition would be. To illustrate the point imagine all possible worlds in which the evolution of human beings takes place. Had Kant been correct in his assumption that the categories of thought were necessary for any possible experience, all rational beings would necessarily come out with the same cognitive schemata in each of these possible worlds. The Darwinians, however, would have to accept that the cognitive schemata and mechanisms could have been different in other possible worlds. All that the reflective mind can make use of is the way the cognitive schemata have developed by chance in the actual world and apply them to the actual appearances.

The Darwinian can say, or rather would aver, that there is an external world since evolution has adapted our thoughts to believe that information received by the senses is not produced by ourselves but has an external origin. We could not have evolved a skeptical disposition toward the origin of this information. We would simply have passed away. So the existence of an external world is an undeniable part of common sense. The next step for the Darwinian is to ensure that substantial ontological claims about the external world may have a cognitive content: This they have if, and only if, the truth of these claims can be justified in virtue of the cognitive resources that evolution has put into our disposal. The cognitive resources we have in our possession include sensations, cognitive mechanisms to process sensory information, purposeful actions as well as a language to express ontological claims. The Darwinian would keep his substantial ontological claims as close as possible to what can be known by the senses. He maintains that there is an external world, but he see no reason to claim that it is objective if one takes objectivity to mean that it has a definite nature in itself, independently and beyond of human cognitive awareness of it. It is reasonable to assume, she argues, that the evolution of our cognitive resources started out with the capacity of organisms to obtain physical signals from the environment in the form of cell receptors to external stimuli. All other mechanisms have adapted around this capacity as various responses such that they have enhanced rather than weakened our capabilities of learning from such stimuli. So for the Darwinian substantial ontological claims concerning the external world can only be justified up to the point where our intellectual powers can still appeal to our sensory experiences. It does not make sense, according to the Darwinian, that something should really exist that we cannot in principle be aware of or with which we cannot in principle interact, and it is a misuse of our intellectual powers to believe that they can grasp a reality to which they cannot possibly be adapted.

6.2 What Makes an Entity Abstract?

All sensory information is about concrete objects and events. The mind can distinguish this information so it experiences separated objects standing out from their environmental background. So we experience not only concrete particulars as being in space and time, but we also take them, rightly or wrongly, to exist in space and time. Thus, ontologically concrete particulars such as objects exist in space and time, or at least in time, and their place therein constitutes their identity conditions. Often philosophers also point to the causal factor. Objects in space and time are able to interact with other objects and therefore have causal powers and liabilities. Space and time may not form the identity conditions of events, since two different physical events can occur at the same space-time point, but following Donald Davidson's suggestion, having the same causes and the same effects may then provide the necessary identity conditions. So for concrete particulars such as physical events, which are in space and time, the same place in a certain chain of causation is what makes them identical. Indeed, the last claim is problematic because of the apparent circularity, but I shall set this issue aside.

In contrast to concrete objects, abstract particulars are considered to be those that do not share any of these identity conditions. Several different ways to understand what it means to be an abstract entity may be adopted. These diverse conceptions rely on the criterion that is used to pick out an entity as abstract rather than concrete. In the current philosophical debate the following suggestions have been proposed:

1. Abstract entities do not exist in space and time.
2. Abstract entities cannot act causally.
3. Abstract entities are logically incapable of existing separate from other things.
4. Abstract entities are those that are introduced as conceptual abstractions by the way of Frege's abstraction principle.

The first of these attempts to portray abstract entities sees them in direct opposition to concrete entities, which are characterized by having essentially spatio-temporal properties. Things like numbers, sets, universals, and propositions can also be ascribed properties, but none of them have spatio-temporal attributes and therefore cannot be identified in terms of a space-time location. Nonetheless, such things can be said to 'exist' as abstract entities, for their existence must be presupposed in order to make statements about numbers, sets, universals, and propositions true.

Closely associated with the view that abstract entities do not exist in space and time is the claim that they are incapable of having any causal influence on anything. It is, if not impossible, at least very difficult to imagine how something that does not exist in space and time could act upon something in space and time. Already in his own time, it was a common objection against Descartes' dualism that it was incapable of explaining how the mind, essentially not extended in space, could causally act on the body, which is essentially spatially extended.

An alternative conception of abstract entities takes them to be incapable of existing independently of other things.[3] We may, following Aristotle, define a "primary substance" (i.e., a 'being' (*ousia*)) as a concrete particular whose existence does not depend for its existence on any other particular. It then follows, by contrast, that a particular whose existence is dependent on other particulars cannot be a 'substance;' it is therefore an abstract entity in the sense under discussion. A Platonist, in contrast to the Aristotelian, would indeed object to such a specification. For him abstract entities are not dependent on concrete particulars; in fact, the latter is what it is only because it participates in the former. But if we grant that abstract particulars are ontologically dependent on concrete things, it seems to imply by necessity that abstract particulars only exist because concrete particulars exist. But how can abstract particulars depend on concrete objects without being caused by them? Well, the whole depends for its existence on the parts; had it not been for the parts there would be no whole, even though the parts do not cause the whole.

Nonetheless, it could be argued that also properties cannot exist on their own—they have to have a substance to be 'in'—all properties by this definition are 'abstract' universals. But neither could a substance exist without its properties. Of course, properties might never exist 'on their own,' but always 'belong' to some 'being' (substance), and therefore when experienced, they are always this or that concrete thing's properties. In contrast, physical entities always exist 'on their own' and never 'in' something else. So, apparently, substances and properties have a very different ontological status vis à vis the abstract versus concrete discussion. In support one may argue that while entities must exist to make statements that refer to them true or false, the same is not true of properties. I may, for example, define the property of 'omniscience' and proceed to argue that it exists nowhere. However, we could make statements about 'omniscience' (aside from its non-existence) that might be true

[3] See, for instance, Lowe (1998), Chap. 10.

or false. For example I might say "Omniscience is desirable." So predication of properties of properties is in a very different situation from predicating properties of entities. But the argument does not hold water. Some statements about non-existing entities, say, unicorns, are true or false by definition. Moreover, properties of properties, or second order properties, are in fact abstract properties, and statements about them may also be true or false internal to a language.

Of course, it may be possible in thought to separate a substance and its properties where ontologically one depends on the other, even if they cannot be separated in reality. An illustration of such a separation would be whenever we think of a particular statue as being divided from the bronze of which it is made. Nevertheless, this view that abstract universals necessarily depend on concrete particulars seems to exclude events from being concrete particulars, since events cannot exist inseparably from those things they involve. The emission of light cannot exist independently of the source that produces it. But events are concrete particulars to the extent that they exist in space and time, they also participate in causal explanations, and sometimes we even identify a concrete object in virtue of a certain event. A sudden flare on the sky, a supernova, may be used to identify the star that once exploded. So, apparently, an entity can be an abstract one in the sense of being existentially dependent upon other entities, but we can still point to it a concrete particular in terms of having a location in space and time. Indeed, this is not satisfactory. The conclusion must be that ontological independence cannot count as a *sufficient* criterion for distinguishing abstract from concrete.

However, I am critical toward using this criterion even as a *necessary* criterion for two reasons. It excludes Platonism by fiat, and thereby it introduces entities that are existentially dependent on concrete objects but in contrast to them are not particulars in space and time. How is this possible? If abstract universals are existentially inseparable from concrete entities, it may because they are identical, which they are not; it may because they are 'coextensive,' which they are not, or it may because the existence of concrete entities causes abstract universals to exist, which they do not. It makes little sense to say that a particular in space and time causes something that is not a particular. Moreover, causation is not a strong enough relation to establish existential inseparability. But I think that (3) is useful as an indication that so-called abstract universals may entirely lack a mind-independent ontological status. Aristotle postulated a certain mental function, “abstraction,” by which the universal is comprehended in the particular. In my view this implies that according to the Platonist interpretation something is an abstract entity if, and only if, it satisfies criteria (1) or (2), whereas in the Aristotelian interpretation something is an abstract universal within a particular object if, and only if, it fulfills criteria like (3) or (4).

The fourth conception goes back to Frege. The idea is that abstract entities are abstracted from concepts following the Fregean abstraction principles. We sometimes refer to concrete objects by using functional expressions of the form “the *F* of *a*.” For instance, we say “The capital of Canada” or “The quotation of Einstein in this paper.” Interestingly enough, Ottawa might not have been picked as the capital of Canada; it might even today have belonged to France, England, or the USA. It is not part of Ottawa’s essence to be the capital of Canada. Similarly, this paper may have omitted any direct quotation; thus, it is not part of the essence of the paper in question to have a quotation of Einstein. The claim is, however, that in those cases where the functional expression “the *F* of *a*” does not refer to any concrete object, say ‘The direction of a line,’ then it is assumed to fulfill Frege’s abstraction principle.

But how do we find out whether “*F*” an abstract sortal or not? For instance, how do we recognize, when saying “The time of the attack on the Twin Towers in New York,” that the moment to which “time” refers is an abstract or a concrete existent?

Obviously, if *F* is existentially dependent upon *a*, what is needed is a criterion of identity that quantifies over a different kind of objects than over those for which it provides a criterion of identity; that is, it must quantify over objects of the same kind as *a* instead of objects of the same kinds as *F*. This is exactly what Frege’s abstraction principle demands. The identity criterion for such cases can be stated in terms of a biconditional, where one side contains an expression of identity between such objects, which are referred to by means of a functional term relating them to items of the kind quantified over, and the other side contains an expression that ascribes a certain property to these items and states an equivalent relation connecting them. Therefore two particular objects, *a* and *b*, are identical if their functional descriptions fulfill the following equation:

(A) $(x)(y)(\mathrm{f}(x) = \mathrm{f}(y)$ if, and only if, $\mathrm{G}(x)$ & $\mathrm{G}(y)$ & $yRx)$.

Thus, applying (A) to time instants, we get something like the following: the time instant of *a* is identical with the time instant of *b* if, and only if, *a* and *b* are events and they coexist. Moreover (A) can similarly be extended to spatial points as well if we replace events with objects and the relation of coincident.

Still we have to specify the features of *R* that make it an equivalent relation. That feature of *R* does not suffice as the abstracting feature owing to the fact that functional expressions like "The capital of Canada" and "The capital of the world's second largest country," both selecting a concrete particular, fit the principle under discussion. We have namely that the capital of *a* is identical with the capital of *b* if, and only if, *a* and *b* are countries and *a* and *b* coincide.

Bob Hale proposes, however, that *R* should meet some further requirements. He argues that

> *F* is an abstract sortal iff, for any *R* that grounds *F*, either (1) *R* cannot hold between spatially located items at all or (2) *R* can hold between things that are spatially, but not temporally, separated.[4]

The first proviso rests on the natural idea that *R* need not exist in space, since the objects over which we quantify can be abstract objects themselves. The second proviso says that in case the quantification runs over concrete objects, *R* cannot connect temporally separated items (because *R* would then not be symmetric). What is important then is how the grounding relation *R* attaches *F* with *a* while it excludes *F* from being a concrete sortal. Concerning *R*, Hale urges:

> *R* grounds *F* iff, for any statement of identity linking *F*-denoting terms, there is some statement to the effect that *R* holds among certain things, the truth of which is (logically) necessary and sufficient for the truth of that statement of *F*-identity.[5]

This characterization of the grounding relation does not secure that (A) always picks out an abstract entity. This can easily be seen in case we

---

[4] Hale (1987), p. 61.

[5] Ibid., p. 59.

substitute "The capital of Canada" and "The capital of the world's second largest country" into (A). The truth of the right side is both logically necessary and sufficient for the truth of the left-hand side.[6] Thus, it seems clear that an abstraction principle of such a simple form as (A) cannot function alone; it only provides a supplement of identity conditions to the conception according to which an abstract entity is existentially dependent and identity dependent upon another entity.

The entity to which *F* refers is one whose existence strongly depends upon *a* and therefore whose identity essentially depends upon *a*.[7] So using the abstraction principle (A), we do not get any further help in determining whether *F* denotes or does not denote an abstract unless we know this in advance: The nature of an entity whose essence it is to be *F* of *a* is an abstract. What we need to know in advance, it seems, is whether *F* is logically incapable of existing independently of *a*.

6.3 Abstract Objects Versus Abstracted Concepts

Common-sense realism does not have room for abstract objects as beings. Plato was extremely aware of this fact when he argued that the ultimate reality consists of permanent forms or ideas that are separated from the world of fluctuating phenomena. What characterizes his forms or ideas is what characterizes abstract objects. They are assumed not to exist in space and time as well as being causally inert. In contrast concrete entities include any entity that exists in space and time and any entity that can be created, interacted with, or destroyed. The classical problem is how one can know that such abstract objects exist if they cannot impinge on our senses. As causally inert they cannot be objects for empirical scrutiny. Plato had few problems concerning the alien nature of abstract objects. He thought that the mind was equipped with a capacity to intuit them. But already Aristotle was skeptic about such a faculty, and much later Hume ruled out any idea that did not originate in sense impressions.

[6] See Burgess & Rosen (1997) for further criticism. See also Lowe (1998), pp. 52–53, where he argues that Hale's proposal is fatally flawed.

[7] See Lowe (1998), pp. 215–216.

What we cannot directly experience but still believe to exist is due to habits and propensities. The Darwinian joins in, arguing that abstract objects, which may include Kant's things-in-themselves, are assumed to exist independently and beyond the grasp of human sensory and intellectual capacities. The metaphysical realist simply lacks any argument demonstrating that biological evolution—besides providing us with the ability to register things in our physical environment—also have furnished us with a faculty, like rational intuition, of being aware of a reality which we have no causally prompted information about. For this reason the metaphysical realist simply postulates the existence of things that are inaccessible to human faculties.

Still the metaphysical realist may argue for the existence of abstract entities because these are thought to be indispensable as truth-makers. For instance, mathematical sentences are true; therefore mathematical objects are necessary for them to be true. Nobody would argue that sentences about concrete entities are true of these objects if the objects mentioned by the sentences do not exist. The same must apply to mathematical sentences; they can be true only if something makes them true, which are therefore assumed to be "mathematical entities." Numbers as abstract entities do not exist in physical space and time but we have knowledge of them even though they do not causally influence our mind. Indeed the evolutionary naturalist, being a metaphysical antirealist, would deny that abstract entities belong to what really exists. Instead, he argued that mathematical objects, like other seemingly abstract entities, exist in virtue of being constructed by us. They are not independently real but of our own creations. So sentences about abstract objects are not true in virtue of a correspondence with some fictitious abstract entities; they are true because we have defined under which conditions such sentences are true and under which they are false.

So the evolutionary naturalist abstains from calling abstractions "real entities," and in doing so her point is that these abstractions do not have any "being" apart from human thought. An abstract object is nothing but an unjustifiably reified concept or mental construction. The term "entity" is the paradigmatic word for saying something has being, or is real on its own. However, Aristotelian *ousia* (beings) were anything that can be the subject of true or false predications; so in that Aristotelian sense the abstractions of science are "beings" since theory allows us to make true or false statements about them. But after Descartes, we distinguish between particular substances, which have formal reality—the kind of reality something has in virtue of the kind of thing it is—and those ideas that have only "objective reality"—they can solely be objects of thought. The distinction between formal existence and objective existence is indeed a scholastic distinction, which is like the newer distinction between objective and subjective. The formal existence is the reality a thing possesses in virtue of being an actual or existent thing and not an existence that it has because it is of a certain kind. When you think of something (i.e., experiencing an idea of it), this has objective existence. It exists as an object of thought. Descartes used this distinction to help him in bridging his ontology and epistemology. The evolutionary naturalist sees the matter in a different light. Abstract objects are objects of thought, and as such they are real with respect to these thoughts, but they have no objective reality in the contemporary sense, as I shall argue in the next chapters. Thus, according to the evolutionary naturalist, *abstract objects are concepts abstracted from sensory information about concrete individuals and then reified as being independently real.*

Let us conclude this section by returning to Carnap's view. In his famous paper "Empiricism, Semantics, and Ontology," Carnap presented what was a very significant challenge for any study of ontology and metaphysics in general. Specifically, Carnap was concerned with abstract objects and the possibility of putting forward substantial ontological claims concerning the real existence of such abstract objects. Non-technically formulated, Carnap's challenge consists in the claim that any assertion about the existence of abstract objects is not an assertion within the language of the assertion but an assertion about that language or more precisely an assertion whose truth or falsehood is fully determined by the rules for the language in which the particular utterance is asserted. It is in this sense that it is postulated that existence claims concerning abstract objects are not substantive ontological claims.

Using a Carnapian terminology, but not his definitions, we may say that an existence claim may be interpreted internally or externally. Internally interpreted existence claims concern phenomena that exist in virtue of the language itself in which they are asserted, but externally interpreted it concerns the actual things that exist independently of the language in which they are asserted. Abstract *concepts* are such language-dependent constructions, whereas abstract *objects* are assumed to be language-independent entities. The task of the evolutionary naturalist is to show that the inference to the existence of abstract objects is unsound and unnecessary; all that can be soundly inferred and that it is necessary to assume is that abstract concepts exist as human constructions, that is, as ideas in human minds. All existence claims that allegedly refer to abstracted entities have no cognitive content, unless they are construed as internal existence claims within that very language.

Abstract concepts are like social constructions such as money, marriages, trade unions, and governments—they exist in virtue of our own thinking and actions. They do not have an ontological status apart from what we make of them. But abstract concepts are also different from social constructions. Where social constructions typically are language based, abstract concepts are biologically based. The first group of entities exists because of intentionally established conventions and the social ontology emerging from it may vary from one culture to another. The second group has its root in our biological adaptation and the ontology assumed by the currently best available theory of organic evolution. If we accept this paradigm we are committed to this ontology.

As providing a prototype ontology concepts appeared in the evolution of cognition as neural blueprints for organization, navigation, and prediction in the organism's adaptation to its environment. Apparently, the first abstracted concepts were those of perceptual sorts, which required a sensory capacity for distinguishing similarities and dissimilarities among the sensory impressions of the organism, a capacity that eventually resulted in the forming a mental response to the similarities among different spatially or temporally distributed kinds. The first level of abstraction is from "concrete individuals," and once concepts for the lowest *infimae species* arise, then these kinds are treated as the "individuals" of higher abstracted "kinds." Much later the reflective mind of philosophers constructed a matching ontology as if the abstracted concepts corresponded to some features of the world over and above the very concrete individual exemplars.

6.4 Why Did Abstracted Concepts Evolve?

Why do we need abstracted concepts but not ontologically self-subsisting abstract objects? There must be an evolutionary explanation of our ability to think of things as something that can be absent from our experiential field and still exist. Likewise, there must be an evolutionary account of why we form concepts abstracted from similar but numerically different things. It has been a huge advantage for our survival that we can imagine that something constitutes a unity in space and time so we can identify two experiences as of the same object when this object has disappeared out of sight in between these two experiences. Such a unity of experience is possible only if we imagine the existence of space and time as providing the continuation of identity of various substances. Similarly, we have also benefitted from being able to consider spatially and temporally separated objects as belonging to the same kind and from classifying properties, e.g., colors, shapes, etc., into the same categories. The ability of mental abstraction from common features among discrete objects in our environment has helped us in identifying dissimilar objects and thereby forming our actions differently toward dissimilar types of objects. The rise of language has even improved this ability in virtue of the inclusion of imperceptible features in the linguistic characterization of invisible objects.

Another type of abstraction stems from our inherited ability to reflect consciously on our own thinking and cognitive practice. The separation of sounds, names, and utterances from what they designate is part of our general cognitive mechanics of separating and identifying objects of thought, an ability that has strong roots in the general conception of the external versus the internal world. Our reflection of our own cognitive practice includes reflection on our thinking in order to learn what is common about our ways of thinking and perhaps improve on what is the most rational method to use in various situations. This led *Homo sapiens* to form civic societies for protection and collaboration and to explore nature for their own prosperity.

We have also benefitted from the ability to think of numbers as not associated with any particular objects. Eventually thinking of numbers was abstracted by reflection from groups of individual objects. Instead, numbers were understood as mental tags that could be used to designate any accidentally chosen objects, and counting was considered to be a numerically ordering of a set of chosen objects as they appeared separated in space and time. Apparently, the potential infinite series of numbers stems from the ability of our imagination to envision that we can always tag a further object in relation to those that have already be named and counted. The capacity of counting helped us to keep track of offspring, groups, games, etc. Reproduction is good for the survival of ones' genes, but for very low reproductive animals like *Homo sapiens*, mating is no guarantee for success unless one has a way of realizing that one can be outnumbered by hostile humans or fierce animals.

So it seems that our disposition to form concepts abstracted from sensory particulars has been selected for two adaptive reasons. It enables us to identify and individuate concrete objects, properties, and relation across different sensory experiences. We are perceptually aware of merely what is present here and now, not of what exists elsewhere or what has been present and is going to be present. Thus, any ability to transcend the information confined to our immediate sensory modalities would make an organism more successful. Moreover, reflective consciousness could not evolve unless an aptitude to identify and classify our own perceptions, imaginations, feelings, emotions, and thoughts had co-evolved. Self-reflection and the ability to classify our intentions help us to plan and act according to those wishes.

The advantage in forming concepts abstracted from regularities among our sensations is quite understandable from an evolutionary point of view. However, it is impossible to explain the evolutionary benefit from assuming that concepts exist as abstract entities independently of the human mind. Considered as abstract objects, such entities could not have had any impact on biological selection. Thus, it does not make sense to claim that abstract objects exist, since we cannot have been adapted to know their existence. We can know only what is of a biological advantage for us to know, but alleged abstract objects cannot be among those things.

7

In Defense of Nominalism

Our everyday world is made up of a countless number of things. In daily discourse we take the existence of these entities for granted because of our awareness of their sensory presentation to us. Thus, the world of experience consists of many different kinds of things that are divided into a plurality of ontologies. Many of these objects either have a proper name or are called by a kind term. In the tradition following Quine many philosophers think of an ontology as an interpretation of a language that makes the language true by assigning a domain of object to the individual variables and let the variables run over a series of properties and relations that satisfy the predicates. Thus, an ontology can be defined as a group of objects that is taken to exist in order to make a language true and that is somehow closely conceptually or functionally connected. If natural, they are assumed to belong to the same family of objects; if artificial, their kinship is constructed. This could be the ontology of hunting or the ontology of shoemaking.

Looking at science, we find that the various sciences do not reduce this number of ontologies but constantly expand it in order to understand newly discovered phenomena under investigation. We have for instance the ontology of anatomy, of physiology, of metabolism, of cell biology,

J. Faye, *Experience and Beyond*,
DOI 10.1007/978-3-319-31077-0_7

of molecular biology, of neuroscience, etc. The practice of science seems to speak against the high hopes of some philosophers and scientists that from a scientific perspective it is always possible to reduce the plurality of ontologies to some fundamental ontology such as physics.

The evolutionary naturalist has no reason to assume that universal reductionism or eliminativism is possible. According to him, physics is as much a result of our cognitive powers as any other scientific discipline. Classical theory reduction, where a set of phenomenological laws is reduced to some more fundamental laws, has been obtained in only a few cases; even if one opts for a more modest reductionism in the form of reductive explanations, where one explains one kind of entity in terms of another kind, such a weak sort of reduction often also fails because we ascribe properties to the whole system that cannot be explained in virtue of their constituents. Thus, as matter of fact, our understanding of most disciplines is such that each constitutes its own ontology. We report our observations in different vocabularies from one science to the next and therefore conceptually classify them differently. So what interests us here is not so much the debate over the reduction of ontologies to a more fundamental set as what makes us classify a group of objects under the same concept. How does this plurality of ontological categories appear in the first place? How do we get to the idea that some particulars are so similar to one another that we take them to be of the same kind?

From an evolutionary perspective, sense organs are adapted to pick up information about concrete objects. Natural selection has produced those creatures that are able to separate and individuate their physical environment into particulars. Particulars have to stand out from the sensory background to give experiential focus and direction to the organism's awareness. In this way many species are able to deal with physical hindrances, finding food, and escaping other hostile creatures. This capacity for sensing concrete particulars also allows many species to identify their mates and to breed and raise offspring. The mechanism of separation and individuation allows orientation, protection, food finding, mate selection, and, if necessary, offspring feeding. Objects appear to have individual properties that not only make it possible for animals to experience these objects as particulars, but also to identify them as the same particulars over time. Based on the ability to identify and remember similarities in perceptual stimuli, these creatures are able to sort concrete particulars into perceptual classes and thereby demonstrate their capacity of learning perceptual concepts.

However, human beings share with all other animals the ability to classify perceptual stimuli into objects of different sorts. Apparently, particulars that have one or more properties in common are perceived as in some way similar. Is it, as realists suggest, just because some particulars share some common properties that make them belong to the same kind, or is it, as the nominalists maintain, just because the concept of a particular kind is constructed by putting all objects that are quite similar into the same category? The simple perceptual ontology of individual kinds of animals is highly restricted in comparison to the many complex ontologies of human beings, because of animals' limited conceptual capacity. But the fact that birds and mammals are able to conceptualize a bit of their environment does not *per se* exclude the possibility that individual objects and properties are what they are because they participate in independently real universals. Nevertheless, nonhuman animals can perceptually categorize sensory stimuli into particulars as well as natural kinds, and we know from evolutionary biology that such a capacity has evolved as the outcome of many selections of accidental variations proved to be useful adaptations to the physical environment. Thus, it seems very unlikely that this capacity at the same time should have been formed in virtue of an evolution caused by the existence of a realm of abstract universals existing separately from concrete particulars.

7.1 Concrete, Artificial, and Nominal Particulars

Evolutionary naturalism divides the world into three main categories. First, there are all of those concrete particulars of which the natural world consists, including of course those particulars whose existence can be experientially or experimentally established. These are the *natural entities*. Some of them have helped determine the selection of every animal's sensory and cognitive capacities. Then we have the *artificial entities,* which are created and designed by human beings and some higher animals for certain purposes. These particulars may comprise physical things like tools, houses, bridges, cars, crops, breeding, and new organisms bred by artificial selection or genetic engineering, but also social institutions such as nations, governments, corporations, universities, wars, marriages, and works of art. Finally, there are what might be call *nominal constructions*. Humans have constructed these in order to understand the concrete particulars. Nominal constructions exist as mental creations, and like artificial particulars they have been designed for certain purposes. Nevertheless, they can be distinguished from artificial entities, as we shall see, because they are not only causally inert by themselves but even our beliefs in them have no causal consequences. They are fictitious, but by imagining them as real nevertheless helps us to structure our thoughts. Among the nominal constructions we find universals, necessities, possible worlds, and numbers.

Now natural entities exist in space and time as concrete particulars, and therefore we are often able to individuate each of them in terms of its spatio-temporal positions because each one cannot occupy more than one place at a time. This invariably holds for material objects as long as we do not grant them the possibility of time travel.[1] Being concrete particulars also associates causal powers with natural and artificial entities. However, if two or more events can occupy the same place at the same time, there may be a problem with this way of individuating physical events. But this is not a problem we shall address here other than pointing to causation as a possible principle of individuation.[2] Turning instead to the artificial particulars we find that many of them have a confined physical existence in space and time that allows us to individuate them accordingly. Being an artificial entity also implies that, though physical, it has been designed to fulfill a certain human purpose. Therefore, some artificial entities can be identified by their function. So you can look at it as either a concrete or an artificial particular depending on how you chose to identify it.

What about the non-physical of the artificial entities such as governments and marriages? In general we cannot directly identify them as having a definite place in space as we can with most physical particulars.

[1] In contrast to physical objects electromagnetic fields, for instance, do not have particular spatio-temporal positions but nevertheless exist in space and time. They also belong to the category of natural entities and as such they can be seen as concrete particulars.

[2] See Davidson (1969).

It seems impossible to argue that a marriage exists in a certain place. A man may be spatially separated from his wife and still be married to her. Nor shall we say in this situation that their marriage is spatially extended, whereas when they are at home together it is spatially unextended. However, a marriage or a government has a precise extension in time because of social actions and rules that define their beginning and end. Usually the existence of this kind of artificial particulars is taken to be a result of a social construction, which means, as John Searle argues, that humans as intentional beings outline the rules or conditions under which something is a marriage or a government.[3] This may well be true. At the same time both a marriage and a government supervene upon human beings' mutual beliefs about certain social and physical facts. These facts may be connected to actions and the physical artifacts of these actions such as signed documents, testimonies, utterances, texts, and paintings. Both the marriage and the government exist in virtue of the institutional relations that exist between two persons or a number of persons, and institutional relations exist as long as the involved partners take them to exist. The existence of institutional relations depends on human beliefs in such facts and their existence is expressed through people's behavior.

Thus, a particular marriage or a particular government can be individuated by specifying the involved people, and a company can be individuated by its articles of association or public registration, etc., all of whom or which exist in space and time. That social institutions can be so individuated and therefore exist as concrete particulars is due to the fact that the shared beliefs in them have causal impact on the believers as well as non-believers. Not all artificial entities exist at a particular place, but since all are designed for behavioral purposes, our beliefs in them constrain our behavior and have important consequences for our actions. If this analysis is correct, it means that our belief in such artificial entities establishes them as concrete particulars because of the causal role these beliefs have for forming our action.[4]

[3] Searle (1995).

[4] Suppose all parties to marriage get amnesia; thus no one believes they are married. Is the marriage then dissolved? It has no causal effect on their behaviour. If also no documents exist, I have no problem in saying yes to the question.

All this seems to be accountable from an evolutionary perspective. Natural selection has given human beings an intentional capacity for forming social behaviors, which is much more advanced than the instinctual behavior already given to us as part of our generic heritage. Our brains coordinate our behavior with our sense organs, and we receive all the information about the physical world on which our behavior depends though them. And the way we are genetically adapted to distinguish particular objects from one another, namely by singling them out in relation to how we sense their position and movements, is still an indispensable part of our nature when we developed new social structures. So the only ability we have to individuate artificial entities is the capacity to relate their existence to something we can observe and act upon. This requires things that have a more permanent physical appearance so that our beliefs about social institutions can build on them.

The ontological analysis changes completely when we discuss nominal constructions. These are similar to artificial entities since they are created for fulfilling a certain aim. But apart from that they are different. They cannot be individuated in relation to physical objects existing in space and time, because they do not exist in space and time. Also, they do not have the capacity to interact with physical objects. Nor do our beliefs in the existence of nominal constructions have any causal influence on the physical world or on other nominal constructions. They are merely inventions. Universals seem to belong to this category.

7.2 Particulars and Universals

I have argued that higher animals behave in ways that makes it plausible to say that they possess concepts in virtue of their ability to distinguish between different kinds of objects.[5] Animals are also capable of distinguishing between various particulars. These abilities are results of natural selection. It had a survival value for those living organisms that were able not only to individuate particulars but also to identify these particulars as of a certain type. Seeing several particulars is one thing; seeing all

[5] See Faye (2014) for a further discussion.

of them as lions is quite another. In the recollection of these animals it became important that they could identify a particular seen earlier to a particular seen later so the visual information in the two cases could be met with the same kind of action. This ability of classification is genetically installed, and it functions as an ability to separate visually invariant particulars from visually variant particulars or, if you like, to distinguish sortal "properties" from accidental "properties." So a living organism with a capacity to form concepts is able to make a mental comparison while ignoring accidental transient "properties" and focusing on the relatively permanent sortal "properties." The latter description can easily be misunderstood. I do not claim that animals are capable of distinguishing "properties" from the particulars themselves. In human beings this separation is due to an intentional abstraction.

That we are able to group particulars into kinds apparently indicates that they share something that makes them similar. Particulars of the same kinds are similar because they have some of the same "properties" in common. This is what the metaphysical realist wishes to believe. Lions are all members of the class of 'lions,' because each individual instantiates in the same universal. Also the colors of their skin are similar because they take part in the same universals. Platonists see universals as abstract objects, which exist independently of particulars (*universalia ante res*). Their existence is necessary for particulars to be what they are as well as necessary truth-makers for statements like "Yellow is the color of this particular." However, the evolutionary naturalist finds such an idea horrendous. How could natural selection bring fitness to organisms with respect to abstract entities? Abstract entities, if they exist at all, do not exist in space and time and are causally inert, which means that they could not have had any effect on the adaptation of organisms, even on our sense organs. One cannot be a Platonist and at the same time believe in natural selection and biological evolution.

Only an environment consisting of concrete entities can have a causal influence on organisms and therefore act as a mechanism of natural selection. By definition abstract universals do not exist in space and time, and what cannot exist in space and time does not have causal powers. But on my arguments so far, universals in the Aristotelian sense (*universalia in rebus*) according to which universals exist within their instances seem not to be ruled out as possible producers of natural selection. Apparently, it is a particular organism that is selected by a particular environment, but it is still possible that this particular organism is selected in virtue of the fact that both the organism and this particular environment instantiate some universals. Universals participating in the individual particulars as their inherent forms and essential properties may explain the fact that natural selection is able to work. Whenever an individual organism is selected, it happens because the actual environment possesses some definite features that the individual fits into. The same features must be found in other parts of the environment if its offspring has to be more fit than any other. Similarly, when evolution produced concept-forming organisms, it may seem possible only if we assumed that those particulars of which these organisms form concepts all instantiate the same universal. How else could the adaptation of genotypes be explained?

Nevertheless, Aristotelian realism is not viable for a couple of reasons. First, if a particular acts as a cause of natural selection in virtue of being an instance of a multi-present (i.e., present simultaneously in a number of particulars) universal, it seems reasonable to assume that it is not only this particular but all particulars instantiating the same multi-present universal that collectively act as the cause—either a particular acts as a cause of natural selection because of its particularity or it acts as such because of its universality. In the former case its particular location in space and time is significant for it to be a cause at all and necessary for it to cause the particular effect in question. In the latter case its particular location in space and time cannot be of any importance. But it is beyond comprehension if a cause becomes a cause only because of a universal that may also numerically be present at the most distance parts of the universe. Second, assume again that the exact same universal exists in many places at the same time. The scarlet here is identical with the scarlet there. But it seems to be a matter of fact that we cannot epistemically separate identical colors from very similar but yet different colors as soon as we cannot use separation spatially as a criterion of individuation. The reason why this scarlet *here* seems identical to that scarlet *there* is not because they are identical, but because these two shades of colors look exactly the same to us. Without space-time separation we have no mind-independent way of individuating multi-present universals. Neither ontologically nor epistemologically do we need universals in order to explain selection and adaptation.

The last claim is supported by the observation that some very few women have an extraordinary color vision. Usually people have a trichomatic vision caused by three different cone cells in our eyes that are susceptible to different wavelengths of light. One kind of cone can distinguish around 100 shades of colors, which in combination yield around one million different gradations (i.e., technical shades, tints, tones, and hues). Thus, an Aristotelian must argue that there are approximately one million color universals. But due to a genetic mutation some women have a fourth kind of cone that in principle makes them tetrachomats allowing them to perceive up to 100 million shades of colors. Only one woman out of 25 who had fourth color vision cones was tested to be tetrachromat. The point is that trichromatic women may judge two shades of colors to be exactly the same, whereas tetrachromatic women may see them as being different. Because their different capabilities for color vision are the result of different neural combinations and ways of processing the information conveyed by the reflected surface light, it does not make sense to argue that either the trichromatic or the tetrachromatic women are able to distinguish the objectively existing shades of colors.

An evolutionary naturalist must be a nominalist of a sort. Nominalists hold that realists erroneously hypostatize mere names into entities, so that the nominalist (at least traditionally) would never want to call these things (if things they be) "entities;" that is the error of the realists. Particulars are the only form of existence she can accept. A nominalist refuses to embrace realism concerning universals on the basis of ontological, semantic, and epistemological arguments. (1) It is ontologically extravagant to claim that universals exist alongside particulars. Universals are explanatorily redundant; (2) semantic features of the language, which it is claimed universals are needed to explain, for example, predicative expressions can be explained in some other manner; (3) even if universals were real, then we could never know if they exist, and what we cannot know gives us no reason to believe in their existence. The evolutionary naturalist denies that human reason can give us insight into a realm of abstract entities or a realm of multi-present universals because human reason is a cognitive disposition adapted to handle sensory experience of particulars.[6] Moreover, the nominalist argues that we can explain the meaning of general names like "scarletness" and the equal meaning of the predicate "is scarlet" in virtue of these terms referring to the class of scarlet particulars.

Nevertheless, the nominalist still has to explain what makes a particular individual "belong" to a certain class. He cannot say that a particular belongs to the class of scarlet things if, and only if, it is scarlet. It is precisely the meaning of "being scarlet" that must be explained without any reference to the property scarletness. Usually, the nominalist claims that a particular fits into the class of scarlet things if, and only if, it belongs to a class of things that resemble each other. The nominal resemblance class is then taken to be either the sum of all particulars that are similar to a paradigmatic particular in the relevant class defining sense or to consist of a 'maximal resemblance class of particulars,' which means to require that all particulars, say those which are scarlet, are included in the class of scarlet things, and that this class is not a subset of a larger class. Such a class is determined in virtue of two randomly chosen members that are similar to each other, and this couple must resemble each other at least as much as they resemble any other particulars that are not members of the class. However, all members of a resemblance class seem to stand in exactly the same relation to any other member of the same class. 'Resemblance' is even a relationship that each member of different resemblance classes has to every other member of its class; thus, each and every member of a resemblance class seems to instantiate a relationship to all the other members of the class that is a universal.

So the nominalist seems to assume what he has to explain. Yes, there are scarlet objects, which are similar, but what makes them similar? Are

[6] Sometimes adaptations originate to serve one purpose but evolve to serve another, as for example feathers were probably originally a means of thermal regulation, but later became an aid to flying. If one supposed there really were a realm of universals, then cognitive abilities originally selected for their ability to deal with sensory particulars could lead to a faculty that evolved to give some kind of rational intuitive insight into these universals, much as Platonists—or for that matter Aristotle and even Descartes—seem to believe. Supposing there really are universals then one could argue that this faculty of rational intuition gave those creatures endowed with it, an advantage in survival and reproduction over those lacking it. However, such a suggestion suffers from the mistake that it considers universals, which are intuitively recognized, as if they were able to interact physically with the organism and thereby to become known to it.

they similar because we see them so, or are they similar because they really have the same property? If two scarlet objects have objectively the same color, then we are back to square one: we are forced into the territory of metaphysical realism, although not necessarily Platonism. A way for the metaphysical realist to meet nominalism could be to suggest, like Aristotle—at least as he is traditionally interpreted—that universals exist only in the particular things. Some scholars have doubted whether this really was Aristotle's view, because if identical universals are located at two places at once, it means that universals in contrast to substances can be at two places at the same time. But if a particular object is taken to be a bundle of properties, it follows that two particular objects instantiating exactly the same universals must be numerically identical even though they seem to occupy two different space points at the same time. However, another interpretation of Aristotle is possible. Talking about the same universal exist at two places does not imply that it is numerically the same universal one has in mind. One may think of only qualitatively the same properties. If this is correct, Aristotle considered universals to be reducible to tropes: that is, particular properties. Whatever Aristotle's view of universals was, tropes are considered to be abstract particulars. So when a trope theorist (tropist) talks about some objects having the same property, she talks about particular properties that resemble each other so much that they are considered qualitatively identical.

The trope theory postulates that the redness of an isolated red car, for instance, exists to the same extent as the car itself and that this particular property cannot be numerically identified with redness of any other red car, even if two cars are similar in having qualitatively identical colors. These two particular colors, regardless of the fact that they may have the same hue, are numerically different. In other words, every car has its own distinct redness, which is a particular, in the same way as the car itself is a particular. Consequently, the predicate "is red" and the general term "redness" do not refer to a definite redness, for example, the redness of this car or the redness of that car. They refer only to the set of resembling tropes, i.e., a class of tropes where two arbitrary elements are similar, at least as much as the couple is similar to any trope that is not a member of the class.[7]

In spite of the fact that the trope theory is often called trope nominalism, it differs from resemblance nominalism by holding that properties exist apart from concrete particulars. The difference between them is that the resemblance nominalist assumes the similarity relation is an external relation between concrete particulars, whereas the trope nominalist thinks of it as an internal relation between properties. Unlike the resemblance nominalist, the trope nominalist can say that the similarity relation is an internal relation in which the similarity is entirely determined by the characteristics of the properties (tropes), which constitute the relata of the relation. Facts concerning the similarity relations among particular things are based on facts concerning the particular's properties. The resemblance nominalist cannot treat the similarity relation in the same way, since he assumes we must consider properties to be universals, whose reality he then denies. According to him, the resemblance exists between concrete things, not between their concrete properties. He denies properties as such. What he claims is that what makes scarlet things scarlet is that they resemble one another.

Both resemblance nominalism and trope nominalism may seem to be able to meet the ontological presuppositions of evolutionary naturalism. Only particulars exist. It is only those toward which our sensory and cognitive abilities are adapted. However, the trope nominalist believes that tropes are abstracts, but it is problematic, as we shall see, how this kind of entity could have had any influence on evolution. Concrete particulars possess materiality; they can reflect or emit light, produce sounds, flavors, and odors, and can be touched. Biological organisms have evolved in response to the world of concrete particulars from which they have evolved. It is impossible to argue that human beings have been adapted to a world of universals with which they cannot causally interact. How

[7] Tropes, or particularized relations, can be understood in two different ways—either they are considered to be ordered triples composed of a relation and two particulars or to be particularized relations regarded as fundamental with no internal parts or structures. They are abstract, because they are not concrete things or substances even though they are spatio-temporally localized. Nevertheless, they are particulars in contrast to universals. Universals arise from such tropes as resemblance classes.

could natural selection work to yield a disposition for grasping universals if abstract entities cannot even in principle influence the fitness of the genes? Resemblance nominalism can certainly explain how we acquire concepts once we have gained the ability to form them. First it needs a working (not necessarily explicit) concept of similarity (internal cognitive 'blueprints' of two or more individuals that are in some respect "similar") in order to group together particulars that are in some respect similar. This has to be in some way pre-empirically "innate," because it must be used before any other concepts derived from a class of similar things can arise in experience. It also needs the concepts of "one" and "many" (or "more than one") before they can build classes from which to abstract concepts. So when living organisms, and in particular our hominid ancestors, gained the ability to reflect consciously on their sensations, then evolution favored those individuals who, with the help of their memory, reflection, and imagination, could imitate the same innate processes of blueprint formation and generalize from seeing single particulars in the past and the present that were in some respect similar to seeing similar particulars in the future. As this capacity evolved, living organisms came to have a general mechanism installed by which they could learn new concepts by abstraction and induction.

Though the evolutionary naturalist is a nominalist, she can easily explain why humans have developed a strong realist attitude toward the existence of universals. Similarity and resemblance signal invariance and permanence, especially if two particulars residing side by side look exactly the same with respect to some property (but not necessarily all) that both have. In many other situations invariance and permanence in our continuous sensations are considered to be the sign that something exists independently of our perception. When we perceive a moving particular to be one and the same over time, we take this particular to exist independently of us. Undoubtedly, such a disposition has high survival value. Therefore, this basic disposition is used whenever somebody faces invariance and resemblance without taking into account that these features may be due to our own cognitive mechanism. This holds especially for all nominal constructions that we have a tendency to reify and project into a realm of independently existing things.

7.3 Conceptualism

Resemblance nominalism eliminates universals, whereas trope nominalism reduces them to abstract properties of particular individuals. The first position denies that properties exist as a class of particular entities existing independently of the objects that are members of the class, but the second position holds that properties do exist independently of the objects that have those properties. We also noticed that this difference between these two positions can be characterized according to how they treat the similarity relation. But up to now we have not attempted to say which one of these two forms of nominalism is supported by an evolutionary approach to metaphysics. Does our biology have anything to say about the pros and cons of each position?

Imagine a lollipop: it is red, sweet, and round. The example stems from C.D. Williams as he introduced the trope theory back in the beginning of the 1950s.[8] So *this-red-now*, *this-sweet-now*, and *this-round-now* are three tropes that are distinct. But how do we distinguish them from one another? It is evident that we cannot appeal to their place in space and time. In general, it seems to be the case that all things have more than one property at a time, regardless of whether we consider big things or very small things. Thus, Keith Campbell was the first to argue that tropes are abstract (because their existence logically depends on the existence of other things) and that the abstractness of tropes is a result of a mental operation.[9] Campbell believes that we get to the lollipop's redness by abstracting it away from the other tropes such as its flavor and shape. This does not imply that the red is a purely mental construction, it may be partly or completely mind-independent (although of course not species-independent), but we need a cognitive act of isolating it from the other tropes that form part of our perception of the whole object. But I would add that such a cognitive operation cannot take place unless we have already formed a concept of redness in virtue of which we can recognize the lollipop as being red. However, this concept of redness is itself a result of mentally isolating the redness tropes of various red objects by

[8] Williams (1953).

[9] Campbell (1981).

carrying out the same kind of cognitive operation while isolating the redness tropes from other kinds of tropes that also belong to these different red objects. We eventually realize that in spite of the fact that red things may seem to be very different, they give rise to a similar experience of seeing a red individual object.

Now, traditionally empiricists distinguished between "general" and "abstract" ideas (or concepts). A "general" idea is allowed. For instance, Berkeley's example is that an idea of triangles in general or of men in general is formed from particular perceptions. Here we simply ignore the characteristics that make one individual in the class different from another. But Berkeley (and Hume) denied that there are any "abstract general ideas," i.e., an idea of a triangle, which is neither large nor small, nor equilateral, nor scalene, nor isosceles, etc. So the idea of, say, being scarlet can be a general idea and can stand for the color of a multitude of individual objects that I can perceive, but he vigorously rejected the abstract general idea of 'color' that is neither red nor blue, nor green, etc., but allegedly 'abstracted' from all the particular colors. Thus, if it is a particular red considered apart from the lollipop or a rose or a fire engine that might have that property, then it is a "general" idea, but there is nothing 'abstract' (in the traditional sense) about a particular shade of red. An abstract idea, in this context, would be an attempt to form an idea of all the various hues of red, what they all have 'in common' abstracted from what makes one hue of red different from another. Nominalists reject that there are any such ideas. We can utter general words and each of them signifies indifferently a plurality of particular things having similar properties. Such words do not denote any abstract concept; their universality stems from how such words are used or function in common language. So, in contrast to a trope nominalist, the resemblance nominalist denies that a general word corresponds to a class of abstract particulars. Apart from that the resemblance nominalist will argue that the concept of tropes is isolated by abstraction from concrete particulars.

Color tropes are only partially mind-independent, because they are the experiential result of the individual organism's dispositional reaction to the light reflection from the surface structure of an external object. The same partially mind-independent status holds for flavor and tastes. Other tropes are taken to be completely mind-independent, such as mass, electric charge, or spin. But any one trope is distinct from any other trope regardless how similar we think they are. So there must be something that makes a scarlet trope more similar to another scarlet trope than to, say, a crimson trope or color tropes more similar to one another than shape tropes, etc. The evolutionary naturalist would say that the capacity to be aware of some resemblance relations has a genetic origin, but others have a reflective origin.

It is an organism's genetic inheritance that determines which colors and shades of colors it is capable of seeing. Let us just assume, for the sake of argument, that the colors we see are mind-independent properties of the objects instead of being a result of our experiential response to light reflection from surface structures. Two scarlet objects may be similar either because they are similar by nature or by our selective adaptation of sorting things into colored classes, as being two particular tropes their resemblance consists in an internal relation fixed completely by the two scarlet tropes or in an external relation fixed by our adaptation to see them as similar.

If we pay closer attention to the internal relation of trope nominalism, it has been recognized in the literature that there might be a problem.[10] As a starting point, the trope theorist must argue that the similarity relation between two tropes is a trope itself. It is just as much an abstract particular as the particular scarlet properties it connects. However, since this particular similarity relation at the same time forms a class of resemblance tropes together with the other particular resemblance relations involving other scarlet tropes, we have a set of similarity relations, R_1, R_2, … R_n,

[10] I still adhere to the criticism that was expressed in the Introduction to Faye, Scheffler, and Urchs (2000) as a reaction to Bacon (1995) who considers universals to be a construction of a set of maximal bundles of tropes: "Similarity turns out to be a troublesome relation, especially when combined with maximality. Somebody's being in love with somebody else, as well as somebody's hating somebody else, seem both to be entirely included in somebody's feeling a strong affection towards somebody else. That excludes at least two of them from being bundles in Bacon's sense, i.e. simple universals. Even a desperate differentiation between 'approximate' and ´precise' similarity doesn't improve the situation. His construction broadly rests upon the assumption of non-actually existing possible worlds. Besides the totality of existent tropes, which form the (actual) world, there are non-existent tropes and possible worlds, formed by (every?) sets of tropes. Without such non-actual worlds, which are motivated mainly by their facilitating the explication of uninstantiated properties, Bacon's trope theory seems to collapse into a rather modest metaphysical construction" (pp. 27–28).

which are all more similar to one another than any particular relation which connects tropes outside of this class, since it is this set of relations that establishes the class of scarlet tropes. The consequence is that we need to assume the existence of another set of similarity relations, Ω_1, Ω_2, … Ω_n, that makes the first set of similarity relations, R_1, R_2, … R_n, more alike in comparison with other resemblance tropes. So trope nominalism seems to presuppose an infinite number of sets of similarity relations or has to introduce a resemblance universal that can stop the regress.

This may be one reason for distinguishing general ideas from abstract ideas. I can say traditional empiricists form a general idea of red by using the word "red" to describe the color of any number of objects of different hues or shades of red, but I cannot 'abstract' from the particular shades of red, that which they all have in common and form an allegedly *abstract* general idea of 'red' that is neither a dark red, nor a bright red, nor a pale, etc., etc. So I can have a general idea of 'relations' that I can use to refer to any particular kind of relation, but I cannot have an *abstract* idea (trope in this context) of relations that is not this relation or that relation or any particular relation but allegedly what all of them have in common.

However, the status of the similarity relation is not the only problem. The trope nominalist considers the similarity relation to be an internal relation, but this implies that the relation exists independently of the human species. Internal relations are an objective part of the concrete world of particulars. But since similarity relations are not themselves material objects, they cannot be the objects of experience in any way other than in virtue of the similar properties of two or more objects. This creates a problem in relation to human evolution, for how can we recognize the existence of such internal relations if we have visual access merely to their alleged relata? In my opinion the trope nominalist has the right to assume that there are resemblance tropes only if she has a species-independent argument for holding that there exist internal relations by which tropes can be connected. Nothing in our knowledge of human evolution supports such an argument.

We know from experiments in cognitive science that conceptual representation and semantic representation are partly different ways of organizing our thoughts. When we are first born our color perception is structured into categories that seem to be the same across all cultures.[11] But once we have learned a language, this partly affects how we categorize our color perceptions. Probably the same holds for other sensory modalities. If this is correct, one may wonder what color trope an object may really have, since the categorization of colors changes partly with respect to one's linguistic community. What we can perceive is that a certain object has a particular color trope, only because it is colored; we cannot specify which particular colors may be internally similar to which. Indeed, according to the trope nominalist, primary qualities such as mass, velocity, angular momentum, charge, etc., are also taken to be tropes, and their similarity relations do not change in relation to a classification. But this still leaves the trope nominalist with the general similarity relation problem unsolved.

The resemblance nominalist does not have to face the same problem. Scarlet things do not resemble one another because they are scarlet; they belong to the class of scarlet things because they perceptually resemble each other. For him the resemblance relation that established a class of scarlet particulars is external with respect to these particulars. An external relation of resemblance cannot exist merely in virtue of its relata since the property of being scarlet is not what makes these particulars similar to one another. The resemblance nominalist normally introduces the existence of a linguistic convention into such a predication. But thereby the resemblance nominalist generates her own problems. Either there must be some external feature that characterizes this class of objects and guarantees that the predication "is scarlet" applies to all these particulars and makes the application stable over time, or, alternatively, the trope theorist can just point to the existence of scarlet tropes, or some external factor such as the human mind is what constructs the resemblance relation. But again the mind could not be adapted to experience various particulars as belonging to a certain type unless our cognitive apparatus was not developed according to the existence of some external entities whose properties we were adapted to experience.

An evolutionary naturalist may therefore argue that predication holds only after language has evolved in a late stage of human cognitive evolution. Before language evolved, similarity relations between concrete

[11] See Franklin et al. (2005) and (2008).

particulars were established in the form of an implicit working concept by the mind's cognitive dispositions of recognition. Our sensory organs respond homogeneously or inconsistently to various stimuli and the brain deals with these stimuli according to the way its sense organs respond. Recognition consists of a uniform cognitive reaction to the homogeneous stimuli. It is in the natural selection of our cognitive functions that we find the explanation of our capacity for comparing objects to see whether they are similar or different. This ability co-evolved with living organisms' ability to form concepts. Likewise, natural selection among living organisms' different responses to their environments explains our ability to see different electromagnetic radiations as colors. We are adapted to see colored things not because these things are really colored, but because we are adapted to register different electromagnetic waves as different perceptual stimuli that we refer to as "colors."

So the evolutionary naturalist ends up subscribing to a form of nominalism that is not very different from classical conceptualism. Particulars are similar because they fall under the general concept of being similar. We cannot say we abstracted this concept from experiences of similar things, because we could not classify these things as 'similar' before we had some notion of similarity. That notion originally arose from the mechanism of a uniform cognitive response to homogenous stimuli, and later as language users we formed a general idea of similarity between various perceptions based on recognition. The capacity of recognizing particulars as "similar" would seem to have obvious advantages for an organism's survival and must have been a very early adaptation in the story of organic evolution. However, universals are nothing but hypostatized concepts (*universalia post res*), formed by reifying certain similarities among concrete entities, thus first guided by natural measures of the sense perceptions and then, when our ancestors acquired a language, by conventional standards designed according to their functions in a conceptual network. Due to our perception, memory, and recognition some abstractions such as color categories may seem more natural than others, but the moment we have the intentional capacity of introducing linguistic conventions we can create as many phony concepts as we like. Thus, their relevance for our thoughts depends on their function in understanding our experience.

8

Space, Time, and Space-Time

The evolution of animals has brought us the ability to form relational classes as birds and higher animals have the capacity of relational concept learning. By the comparison of sensory stimuli they can learned that one object is either the same as or different from another object because they already have some kind of pre-programming for "perceiving" sameness. It is also true that relational concept learning covers spatial and temporal concept relations. Rodents can learn to run a maze, guillemots navigate back to their nesting site, and hunting cheetahs react to the spatial changes of their prey. In animal cognition scientists distinguish between *beacon homing* and *piloting* in which a bird or mammal either moves toward a beacon or a location defined by its spatial relations to perceived, localized clues. Other evidence of animal navigation shows that geometrical properties of an area are coded in some animal's brains as visual relations defining a represented spatial frame of reference with respect to their bodies. This *geometric module* may work together with *dead reckoning* where "internal movement cues (provided primarily by the vestibular system) allow the animal to integrate its position in space relative to a starting point." It is even more interesting that animals can learn by abstraction to use these relations to recognize similar spatial

J. Faye, *Experience and Beyond*,
DOI 10.1007/978-3-319-31077-0_8

patterns from experiencing particular spatial relations among concrete objects.[1]

The same holds for temporal conceptual relations. Biological clocks, circadian rhythms, and model periodical events in birds and mammals' environments allow organisms to grasp timing intervals by linking recurring events to their circadian clocks.[2] Thus, it seems reasonable to believe that our thinking of space and time is grounded in our innate cognitive capacity to form conceptual presentations of spatial and temporal *relations* among individuals based on our bodily, tactile, and visual experiences. Does this imply that space and time do not exist as concrete entities "outside" of our biological organisms, and if they do not, in what sense and extent do their existence rely on our cognitive system?

Traditionally theories about the nature of space and time come in two versions. Some philosophers regard space and time to be substantial in the sense that they consider space-time points to be fundamental entities in their own right independently of their relations to anything else in the universe; others take space and time to refer to relational facts by (somehow) constructing points and moments out of objects and events. In spite of their fundamental disagreements, substantivalists and relationists share a common view: They regard space and time descriptions as referring to concrete particulars. Hence Quine's famous dictum "no entity without identity" should apply to

---

[1] See Brown (2006). In a series of experiments with rats using a 5 × 5 pole box, Brown was able to demonstrate that their choices can be controlled by a spatial pattern among goal locations that does not correspond to any perceptual cues. The evidence for spatial pattern learning is reviewed, and some possible mechanisms are discussed. He concludes: "In order to abstract the spatial relations among goal locations, given that the goal locations change unpredictably in allocentric space, rats must somehow be perceiving the spatial relations among the baited poles found during individual trials. Two possible mechanisms for doing so can be distinguished. First, a *working memory* system could be used to code the allocentric location of poles previously discovered during a trial. The spatial relations among those locations could then be determined on the basis of working memories for their locations. The abstracted spatial relations among baited locations would then be coded in a more permanent memory system. According to this view, the process of spatial pattern learning is analogous to concept learning in that the spatial relations are abstracted from particular exemplars of baited pole locations experienced over trials.

Alternatively, a *dead reckoning* system could be used that integrates the distance and direction from each baited pole discovered to the next. According to this view, rats need not code the locations of particular baited poles during the trial. Instead, their spatial relationship is coded directly in terms of the vector provided by dead reckoning as the rat moves in the pole box and chooses poles. A new vector is initiated each time the rat discovers a baited pole. The resulting set of vectors specifying the relations among each pair of poles forming the pattern constitutes the learned spatial pattern."

[2] See, for instance, Church (2002) and (2003).

space and time loci as well. Supposing there are such things as concrete particulars, we must be able to point to some determinate identity conditions of space and time points that would allow us to regard them as concrete particulars. In fact, most philosophers just take for granted that space and time are concrete entities; they tacitly presume that appropriate identity conditions exist and that it is rather unproblematic to specify what these are. Yet these assumptions are in fact problematic—can we provide a better analysis?

In deliberating about absolute and relational theories of space and time, John Earman points to the serious difficulties concerning identity and individuation any theory of space-time points must confront. After discussing various metaphysical accounts of predication, he makes the following remarks:

> One could try to escape these difficulties by saying of space-time points what has been said of the natural numbers, namely, that they are abstract rather than concrete objects in that they are to be identified with an order type. But this escape route robs space-time points of much of their substantiality and thus renders obscure the meaning of physical determinism understood, as the substantivalist would have it, as a doctrine about the uniqueness of the unfolding of events at space-time locations.[3]

Earman does not go further into this suggestion because, as he observes, it departs too strongly from the substantivalist core assumptions. However, I shall lay out a view according to which our concept of space-time points is an abstraction drawn from what is given empirically.

First I shall review some of the difficulties, which Earman mentions, in the light of recent discussions on the identity and individuality of space and time points. The conclusion of this discussion is that space and time points should be categorized as abstracted particulars.[4] Apparently, Leibniz meant something similar in his correspondence with Clarke when he pointed out that space and time are not fully real but are 'ideals.'[5] Space, I submit, refers to the ordered set of the loci of all bodies extended over all times, and time

[3] Earman (1989), p. 199.

[4] In an earlier paper (Faye 2006b), I argued that time is an abstract entity but kept a door open for the concreteness of space. Also I counted Leibniz as a proponent of space and time as concretes because I took him for being a reductionist by heart. Now, having reconsidered, I must admit that this remark may be too hasty.

[5] Indeed, 'ideal' have several meanings. By using 'ideal' in contrast to 'real,' Leibniz seems to think of space and time as something whose existence (partly) depends on the mind.

designates the ordered set of all changes of these bodies with a determinate beginning and end. I believe that this position has some very important explanatory advantages and that it may even provide a satisfactory solution to the debate between the relationists and the substantivalists. I shall present some arguments to the effect that the points of space and time, or space-time, to which the descriptions of physical processes refer, should be considered as reified abstractions. By this I mean that space-time points are conceptual constructions—manmade artifacts whose role is to help us represent the world by means of identifying and individuating concrete particular objects.

My suggestion is that space-time is an abstraction whose structure is constructed by the human mind from our spatial and temporal comprehension of individual things and events. Indeed, we need to say how much of our comprehension is purely 'perceptual' and how much represents a post-perceptual judgment limited to human beings. Where are the animals in all of this? Surely they also make spatio-temporal judgments. I doubt that animals "live in" space and time any differently than we do. So the question is how we get from the concrete empirically "given" space and time of animal experience to the "abstraction" of the physicists' space-time. What is "given" in experience is, by definition, always "concrete," a particular this or that "with determinate properties." From this evolutionary basis, human thought derives abstractions by generalizing or leaving out various properties of the originally empirically given. So if spatio-temporal comprehension is an adaptive selected experience, then that adaptation happened, evolutionarily speaking, well before humans ever come on the scene. As far as at least the "higher" animals and "us" are concerned, we are all in the same boat with respect to experiencing space and time, and it would take a lot of fancy philosophical dancing (Hopis notwithstanding) to try to show that animals do not in fact experience a world extended through space evolving over time in a manner at least very similar to the way humans do.

Empirically we experience both spatial regions and temporal succession, and in human experience we can directly distinguish the one from the other. But we cannot experience spatial points or temporal instants. Both points and instants have the rather abstract, non-empirical quality of being dimensionless, and a spatio-temporal continuum requires an infinite number of points. Thus, philosophers and physicists loosen the physical concept of space and time from our spatio-temporal experience by arguing that space-time points stand for objective features of an independent world.

In contrast, I take substantive claims about space-time to be claims about reified concepts that have evolved for cognitive reasons as tools for individuating and tracking concrete particulars and whose identity therefore does not fulfill the normal determinate identity condition of concrete objects.

8.1 The Existence of Space

For everyday consciousness *empirical* space and time are perceived as "real," but those realities are very different from the physicists' construction of the *physical* space and time. Empirical space is perceived in the form of location or extension of material objects, whereas we experience empirical time in the form of change and duration of physical processes. Whether in physics we think of space as being absolute or relational, either view takes it for granted that spatial points exist independently of any perceiver's experience of location and extension. The absolutist, in being a substantivalist, believes that spatial points and relations exist over and above what is located in them and that these points have intrinsic relations to one another. The relationist, in contrast, argues that spatial points and relations are nothing by themselves, since they are reducible to relations between things that are said to occupy them. For the sake of terminology I distinguish between "points" and "instants" as geometrical indications of *empirical* locations and changes, and "Space" and "Time" of which all these indications are taken to be *proper geometrical* parts. We can perceive locations and changes, but neither points and instants nor Space and Time; the concepts of Space and Time are mathematical constructions based on spatial and temporal indications such as abstract points and instants.

Consequently, we shall differentiate between (1) *empirical* space and time as an indication of the relations between what can be perceived here and now; (2) our *ordinary* notion of space and time that functions as an extension of empirical space and time based on our memory and expectation to include unperceived things and events; and finally (3) the *physical* notion of space and time, reflecting some *intrinsic* geometrical or mathematical structures of space and time as described by our best physical theories and referred to as *Space, Time,* or *Space-time*. The problem we have to solve is whether space and time in their physical sense can be considered invisible but real entities or whether they should be regarded as abstractions.

Historically, the two characterizations of Space and Time may not be true of the arch contestants of substantivalism and relationism, respectively. Newton denied that "absolute true and mathematical" Space and Time are real substances, nor are they accidents. He seems to have taken over Pierre Gassendi's view that Space and Time are of a third kind, claiming that Space and Time are preconditions of substance. Before Newton, Gassendi argued that Time flows uniformly regardless of any motion and that Space is uniformly extended irrespectively of the bodies it may contain.[6] Newton associated Space and Time with modes of existence because of his assumption of God as the necessary being who is substantially omnipresent and eternal. Nonetheless, he claimed: "Although space may be empty of body, nevertheless it is not itself a void; and *something* is there because spaces are there, though nothing more than that."[7] He also emphasized that space is distinct from body and that bodies fill the space where they are located. So Newton seems to be as close to being a substantivalist as one can be, especially if one brackets his belief in God and considers Space to be a geometrical substance that can exist empty of any material substances. Similarly, Leibniz was less of a hard-core relationist than was Descartes. In his correspondence with Clarke, he explicitly said that Space and Time are 'ideals,' having no full reality. Space "being neither a substance, nor an accident, it must be a mere ideal thing, the consideration of which is nevertheless useful."[8] This is interesting because it indicates that Leibniz saw Space and Time as geometrical abstractions rather than as the aggregate of existing spatial and temporal relations among all material objects.

If spatial points are concrete entities, it must be possible to specify their identity conditions in a way showing that they are concrete entities. Spatial points are in Space, and being in Space is a common criterion of being a concrete entity. However, spatial points cannot exist independently of Space itself. Spatial points are intrinsically featureless; they lack any internal features for differentiation among themselves. Being parts of Space, they have, by necessity, the same nature of identity as Space itself in terms of being concrete

[6] See Gassendi (1971), p. 383 ff.

[7] Newton (1962), p. 138.

[8] Alexander (1956), p. 71.

or abstract. Space is not just the mereological sum of its parts, even though spatial points may seem to be absolutely the same all the way down, because Space, taken to be a substance, contains an absolute metric that cannot emerge from a collection of individual parts. Rather, the individuality of the spatial points comes from the structure of Space itself. Bearing witness to this claim, Newton said: "The parts of duration and space are only understood to be the same as they really are, because of their mutual order and position, nor do they have any hint of individuality apart from that order and position which consequently cannot be altered."[9] A spatial point depends for its existence upon Space, and consequently its identity depends on the identity of Space. Thus, spatial points are concrete particulars if, and only if, Space itself is a concrete particular. But in order for Space to be concrete it must exist in Space. However, Space itself cannot be in Space, because that would make Space a part of Space; thus its identity would depend on this further Space, etc. Therefore spatial points cannot be concrete entities.

The causal criterion of concreteness also does not apply to Space. Although it has been held that Newton considered absolute Space to be a cause of the inertial forces, there is no textual evidence for such an interpretation, and it seems more accurate to say that Newton believed that absolute Space merely acts as a frame of reference and that acceleration by itself gives rise to the inertial forces. The relationist, however, hopes to account for the distinction between relative motion and "real" accelerated motion not in terms of absolute Space, or any other object in relation to which the motion is relative, but in terms of the causes of the motion.

Now, we may find that the locations of things can be defined in terms of functional expression such as "The location of Montreal" is the same as "The location of the largest city of Canada," where the identity conditions of locations is dependent on things occupying them and the spatial relations. At first glance it seems possible to identify locations quite independently of the physical things, which may occupy those locations, in virtue of geometrical indications alone:

- $(x)(y)((\text{loc}x = \text{loc}y)$ if, and only if, point(x) & point(y) & x coincides with $y)$.

[9] Newton (1962), p. 136.

But we have just learned that the individuality of spatial points depends on the order of Space itself; hence, if Space is not a concrete object, neither can spatial points be. Furthermore, we should notice that the relation "coincides with" is reflexive, symmetric, and transitive as required by the abstraction principle. Locations can therefore be pointed out in relation to concrete particulars and their mutual spatial relations.

The proper identity condition for locations is then expressed by a proposition, which grounds the term 'location' in the coincident relation between things or other concrete particulars:

- $(x)(y)((\text{loc}x = \text{loc}y)$ if, and only if, thing(x) & thing(y) & x coincides with y).

Neat as the statement seems, it is nonetheless obvious that it negates the existence of empty space. Avoiding any animosity of the void (between separated things), we must allow a modal formulation like the following:

- $(x)(y)((\text{loc}x = \text{loc}y)$ if, and only if, thing(x) & thing(y) & 1) x coincides with y, or 2) in case y and y did not exist, then if they had existed, x would have coincided with y whenever y would have coincided with x, and vice versa).

This illustrates that locations are actually distinct from physical things but still are logically incapable of existing separately from physical things as such.

8.2 The Existence of Time

Aristotle said that time is not change, but the measure of change, or rather "that in respect of which change is numerable."[10] Thus, a time is our abstract indication of change. This suggestion was perhaps not such a bad proposal. Change is something we can perceive. Together with location and extension, change and motion are what we can immediately see

[10] *Physics*, 219^b2.

with the naked eye, whereas apparently we are able to recognize space and time only indirectly with the help of celestial motion of objects such as the Sun or mechanical clocks. Thus, awareness of time logically presupposes awareness of at least change. But there is more to Aristotle's suggestion than epistemological priority. Also time *is* nothing but a measure of motion. Given this interpretation, motion is not only semantically prior but ontologically prior to time as well. The existence of motion and change precedes the existence of time. His ontology of time thus comes close to our everyday experience of temporality.

This also explains why Aristotle seems to deny the existence of temporal instants. He says in connection with Zeno's paradoxes: Zeno's conclusion "follows from the assumption that time is composed of moments: if this assumption is not granted, the conclusion will not follow."[11] What Aristotle probably had in mind was something like this: Since Time, as a geometrical measure of change, is supposed to be continuous then each period of time must contain an infinite number of instants. But, he assumed that nothing can be *actually* infinite, but only *potentially* infinite. Numbers are in this way infinite in so far as there is no limit built into the process of counting. Likewise we can divide a length or a period of time in as many points or instants as we want, there is no limit to such divisions, but the divisions do not exist independently of the one who makes them. Hence, the potentially infinite divisibility does not imply the existence of actually infinite divisibility, and therefore spatial points and temporal instants do not exist independently of us. Although Aristotle did not explicitly say so, his view is not so far from saying that points and moments (space and times) are not concrete entities, but abstract ones, since they are being the product of the converging limit of our cognitive ability to divide things up into smaller and smaller regions and intervals.

Following up on Aristotle, we may say that Space and Time cannot exist as a measure of motion unless things in motion exist prior to the numbering. Time exists only if change and motion exist. It is impossible for Time to exist in case there is no change or motion. Thus, we see here an exemplification of the conception of abstractness according

[11] *Physics* 239^b30–3.

to which existential dependence marks what it means to be an abstract entity; Time ontologically depends on things in motion or things that undergo change. Moments are abstracted from varying things but do not exist independently of the concrete particulars from which they are abstracted. In contrast, substantivalism—as we find it in Newton's notion of absolute Space and Time—takes moments to be ontologically prior to those physical events that may occupy them; Time exists as an independent entity, whereas relationism regards moments to be identical to physical events or their existence to be somehow parasitic on things and processes. Both views consider Time to be a concrete particular. The first view captures Time as a substance, the second view as a non-substance. This means that it must be possible to specify some identity criteria that show that Time is a concrete particular. But what are they?

Temporal instants seem to be concrete particulars existing in Time, because they stand in temporal relations to other times, and we seem to have no problems of specifying identity criteria for such moments. We say:

- $(t)(t^*)((t = t^*)$ if, and only if, moment(t) & moment(t^*) & t is simultaneous with $t^*)$.

But moments cannot exist independently of Time itself; they are parts of Time, and as parts of Time they must have the same nature of identity as Time itself in the form of either being concrete or being abstract. A moment, i.e., temporal instant, depends for its existence upon Time, which implies that the identity of a temporal instant depends on the identity of Time. Therefore, moments are concrete particulars only if Time itself is a concrete particular.

Assume that Time is a substance. Time should then, like any other physical substance, exist in Space and Time. But Time does not exist in Time, whereas Space may be said to exist in Time; thus, Space and Time cannot determine the identity of Time. Hence, Time cannot be an individual substance. Assume, in contrast, that Time is not a substance because all talk about temporal instants and relations can be reduced to talks about events and causal relations. This requires that we can set up identity criteria of events, which avoid any reference to space and time. Here Davidson's attempt to specify determinate identity criteria of events in terms of causation comes to mind as the only serious suggestion, claiming that:

- $(x)(y)((x = y)$ if, and only, if event(x) & event(y) & x and y cause and are caused by the same events).

Unfortunately, this criterion has rightly and often been charged as being circular.[12] Thus, the conclusion seems to be inescapable. Time cannot be a concrete particular.

In contrast, I propose that Time is an abstraction in the sense that we have a constructed geometrical language to be able to reflect about collections or sets of moments and thus of concrete changes. Time denotes an ordered set of all moments in the world. This suggestion is supported by the above conceptions of abstractness. Time itself does not exist in Space and Time. Again, Time does not have any causal influence on concrete substances because, if it had had such an influence, then each and every particular event would be causally overdetermined by causally prior events and by the definite moment at which the event takes place since both the causally prior events and the moment in question would be causally sufficient for it. Moments are therefore causally superfluous. Moreover, if we think of two events that are causally connected so that the cause is not only causally sufficient, but also causally necessary for the event, there is no room for causally active moments.

We should notice that Time, like events, is logically incapable of existing separate from particular substances. Events, however, in contrast to Time, do exist in empirical space and time, and thus we shall leave aside that they may be abstract in some other sense. Time cannot exist without changing things; nevertheless, we can, of course, separate time instants from changes in thought.

Finally, the concept of a temporal instant fulfills the principle of abstraction. We can assign a time instant to an event in terms of a functional expression and thus express the identity of moments in terms of identity of events. We say, for instance, the time of the Big Bang, the time of the supernova, and the time of the solar eclipse. These functional expressions meet the abstraction principles.

---

[12] Faye (1989), pp. 153–160.

- $(x)(y)((\text{inst}x = \text{inst}y)$ if, and only if, event(x) & event(y) & x coexists with y in relation to a frame S).

It says that the moment of event x and the moment of event y are identical if, and only if, events x and y coexist. The relation 'coexistence' is indeed reflexive, symmetric, and transitive in any given inertial frame, and it grounds the abstract sortal term 'instant' such that the understanding of instants or moments presupposes an understanding of events and changes.

8.3 Space-Time Substantivalism

Up to now we have mainly considered whether Space and Time are concrete or abstract entities in a metaphysical context. Now it is time to continue our analysis to look at the question from the point of view of what physics has to tell us.

In modern physics Space and Time "merge" into a single dynamic entity called Space-time in the sense that they cannot be unambiguously *distinguished* independently of reference frame. It is sometimes assumed that this entity, according to the field equations of the general theory of relativity, is causally efficacious in the sense that Space-time causes the distribution of matter and energy in the universe, which in return affects the curvature of Space-time. This assumption of mutual influence requires that Space-time is a concrete entity that is able to undergo changes that effect or are affected by changes in the matter and energy distribution. However changes, by definition, can take place only in something that persists through these changes. If one accepts this metaphysical principle, it leads to the conclusion that Space-time should be treated as an object, or rather a substance, which forms the persistent ontological ground for any change. Therefore, the assumption that Space-time can be causally influenced by the distribution of matter and energy, or can causally influence that distribution, presupposes substantivalism of some sort. Space-time is a real substance undergoing changes but that exists independently of those processes occurring within Space-time.

Indeed, Space-time substantivalism constitutes a serious threat to the claim that Space and Time are abstractions, because it must treat space-time points as concrete particulars. Proponents of substantivalism point out that the general theory of relativity quantifies over space-time points, and as true followers of Quine they take this as a reason for believing in the existence of space-time points. Therefore we need to take a closer look at this view.

In a co-written paper John Earman and John Norton define 'substantivalism' as the claim that Space-time has an identity independently of the fields contained in it. They emphasize that the equations describing these fields "are simply not sufficiently strong to determine uniquely all the spatio-temporal properties to which the substantivalist is committed."[13] This captures the standard view that a substance is something that is self-subsistent; thus, a substance can be defined as a particular whose identity does not depend on any other particular and whose existence therefore does not depend on it.

Before we proceed an important distinction should be made between *manifold substantivalism* and *metric substantivalism*. The first type forms a kind of minimal view, according to which Space-time consists of a topological manifold of points, and the metric field is then attached as an externally defined field, whereas the second type includes the metric field as an intrinsic part of the Space-time container itself.

Earman and Norton identify Space-time with space-time points. As they say: "Thus we look upon the bare manifold—the 'container' of these fields—as spacetime."[14] The bare manifold consists of space-time points, whereas the fields form the metrical structure of Space-time, which is added to the manifold as a thing in it. Their motivation for separating the bare manifold as the Space-time container and the metric fields as the contained is that the metric fields carry energy and momentum that can be converted into other forms of energy and heat.

Manifold substantivalism assumes space-time points are real, but it is entirely unclear what their identity conditions are. It has been noted before that according to Newton the geometrical parts of absolute Space and Time are intrinsically identical to one another and can only be differentiated by their mutual intrinsic order.[15] But this move is foreclosed to the manifold substantivalist. The identity conditions for space-time points cannot involve the metrical structure because of the way manifold substantivalism has been defined. It is assumed that Space-time is noth-

[13] Earman & Norton (1987), p 516.

[14] Ibid., pp. 518–519.

[15] See also Maudlin (1989), p. 86.

ing over and above space-time points in the sense that the identity of the Space-time manifold is dependent on the identity of space-time points. The consequence is that Space-time is a real concrete particular if, and only if, space-time points have an identity independently of any entity that may occupy these loci.

Nevertheless, it appears reasonable to say that Space-time is not a composite substance because the whole does not distinguish itself from the parts. Space-time is indefinitely divisible into other particulars of the same kind, but how can we distinguish between these parts in such a way that the distinction represents a real difference? Establishing determinate identity conditions, which make Space-time a concrete entity, is a serious problem for manifold substantivalism. The points of the manifold are pure abstract individuals, bare mathematical particulars, which do not have any structure or properties in virtue of themselves. How can we know whether or not these self-subsistent mathematical objects can be represented by physical space-time locations (or events)?

The manifold substantivalist seems to have two possibilities for formulating determinate identity conditions of space-time points. She can either follow the mathematical road or take the physical one. It is possible for her to regard physical points as names of geometrical points, or she can regard some form of metric structure as inherent in Space-time (because we are able to talk about a universe free of matter and energy).

Following the first path, the manifold substantivalist does not collide with the concrete structure of the world, but nonetheless the road is not passable. It is impossible to see how physical locations can act as names for geometrical space-time points, unless we already possess independent physical criteria of individuating particular space-time points. A name refers to what it names. Assuming space-time points to be something considered to be "real," the named has identity conditions distinct from being so named, i.e., that they are mutually independent. The manifold substantivalist, however, is unable to point to what these conditions are with respect to physical space-time locations. Mathematical points are all we have, and they have no intrinsic features that individuate them from each other.

As we have seen, individual space-time points can be defined only relatively within a particular relational structure, and their *only* identity is given in virtue of their position in this structure. We may indeed assign coordinates to the manifold, but in a pure differential manifold each and every possible form of coordinates is arbitrary, and the manifold is invariant with respect to the choice of a particular coordinate system. Only by adding a structure is it possible to change the situation, but then we are no longer confronting a bare manifold.

Choosing the second road, the manifold substantivalist may locate the identity conditions of space-time points in the metric structure of the physical state of the universe (*versus* Earman and Norton). In this way he may attempt to uphold a view of Space-time as a concrete entity. If Space-time is taken to be represented by a manifold of geometrical points on which we define a metric field, then the set of physical space-time points is individuated by their metric properties as they are defined by our best space-time theories.

The general theory of relativity identifies the metric field with the gravitational field, and therefore it carries momentum and energy. Norton describes the problem thus:

> This energy and momentum is freely interchanged with other matter fields in space-times. It is the source of the huge quantities of energy released as radiation and heat in stellar collapse, for example. To carry energy and momentum is a natural distinguishing characteristic of matter contained within space-time. So the metric field of general relativity seems to defy easy characterization. We would like it to be exclusively part of space-time the container, or exclusively part of matter the contained. Yet is seems to be part of both.[16]

Indeed, if the energy and momentum of the gravitational field can be converted into radiation and heat, and vice versa, in connection with the formation of black holes, and this field also characterized the metric properties of Space-time, how can Space-time exist independently of what is going on in it? Since the identity of space-time points logically depends on their metrical structure, they are incapable of existing without this structure.

The manifold substantivalist may respond by pointing out that Einstein's field equations connect the intrinsic structure of Space-time with the distribution of matter and energy such that the metric field, in the form of the gravitational mass-energy field, and the matter field stand in a causal

[16] Norton (2004).

relationship. Thus, if Space-time has no momentum and energy, it would be impossible to see how they could interact with matter. Moreover, we can have a causal relation only if the relata are logically distinct from one another, i.e., in case the relata have mutually independent identity conditions. Thus, if Space-time and stars and galaxies were separate entities, then their mutual causal interactions would constitute the proof that they are concrete particulars. But the argument, as it stands, is not without problems.

I sympathize with Lawrence Sklar's warning over interpreting the field equations as the non-gravitational mass-energy causing modifications of space-time since "the possible distribution of mass-energy throughout a space-time depends upon the intrinsic geometry of that spacetime."[17] Apparently, what he wants to emphasize is that the matter field is spatially and temporally distributed. Thus, it cannot gain the necessary ontological independence of the metric field necessary in order to have the separate existence required for having causal efficiency. Instead, Sklar maintains that the field equation should be interpreted as a law of coexistence:

> The equation tells us that given *both* a certain intrinsic geometry for space-time and a specification of the distribution of mass-energy throughout this spacetime, the joint description is the description of a general-relativistically possible world only if the two descriptions jointly obey the field equation.[18]

Such a law-like constraint on the two descriptions robs the substantivalist of the causal argument for the independent reality of Space-time and for the matter field being concrete, independent particulars.

Where does this line of reasoning lead? It seems that manifold substantivalism either is forced into admitting Space-time is an abstract mathematical entity (since space-time points become abstract particulars) or collapses into a form of relationism where Space-time as such is claimed to be identical with the fields of gravitation-cum-matter. In the latter case the metric field is defined in terms of the gravitational fields, whereas the space-time points are defined in terms of the mass-energy fields. So, if one wants to sustain a claim that Space-time is a concrete substance, manifold substantivalism seems not to be a viable metaphysical possibility.

---

[17] Sklar (1974), p. 75.

[18] Ibid., p 75.

In the debate over *manifold substantivalism*, according to which Space-time is represented by a manifold of points and a metric field is added to this manifold, one argument appears to be more prominent than any other: Einstein's hole argument. If we first label the space-time points in the manifold and then spread the metric fields over it in different ways, as the general covariance of the field equations allow us to do, the different results will describe physically different situations. Thus, a hole in the Space-time manifold would allow different solutions for the metric field inside the hole without changing the boundary conditions. Einstein found his solution in point-coincidence, or in a modern formulation, diffeomorphism invariance, which Earman and Norton dubbed the *Leibniz equivalence* as it represents in their view the space-time analogy of the traditional *Leibniz shift*.[19] Leibniz equivalence is hence the assertion that a single physical situation (a gravitational field) can be represented by an equivalence class of diffeomorphic models.

The hole argument appears to show that a substantivalist interpretation of Space-time requires ascribing a surplus of properties to Space-time, which cannot be determined by observation or the laws of the relevant space-time theories. The substantivalist must concede that matter fields, which after a diffeomorphic transformation go through such a hole in the space-time manifold, are not determined by the metric fields and the matter fields outside the hole. Nevertheless, the manifold substantivalist, who wants to save determinism, also holds that there have to be physical differences between the possible trajectories that a galaxy may take inside the hole. Earman and Norton take this to be a most unwelcome consequence of space-time substantivalism and for this reason are ready to give up manifold substantivalism as such.[20]

Attempts to avoid such a conclusion by adding further structure to the manifold can, at least in some important cases, be met by alternative versions of the hole argument.[21] If manifold substantivalism has to be abandoned, Earman sees three ways to uphold substantivalism with respect to space-time points. One may adopt a structural role theory of identity of space-time points (which I shall return to below in the form of *sophisticated substantivalism*), one may argue that metrical properties are

[19] Earman and Norton (1987)

[20] See Earman (1989), Ch. 9.

[21] See Norton (1988).

essential to space-time points,[22]or one may introduce counterpart theory to space-time models.[23] But in conclusion he finds that that "our initial survey of the possibilities was not encouraging."[24]

The central claim of *metric substantivalism*, according to Tim Maudlin, is that "Physical space-time regions cannot exist without, and maintain no identity apart from, the particular spatio-temporal relations which obtain between them."[25] Thus, the identity conditions of space-time points are determined by the intrinsic order among them. A few pages later he states that space-time and metric are connected by necessity: "Since space-time has its spatio-temporal features essentially (cf. Newton above), the metric is essential to it and matter fields not."[26] The proponent of the *metric substantivalism*, in contrast to the manifold substantivalist, welcomes the idea that Space-time carries energy in the form of its metrical structure, because it puts Space-time on a par with other substances.[27]

In the general theory of relativity, the metric field is associated with the gravitational field because of the proportionality of gravitational and inertial mass, so that gravitation and accelerated coordinate systems can be considered physically equivalent. Einstein spoke about this association in various terms: The gravitational field is said to either *influence* (or determine) or *define* the metrical properties of Space-time.[28] But holding that the gravitational field *defines* the metric structure of Space-time, it must be a conceptual feature of the universe and not just accidental that gravitational and inertial mass is proportional. This indicates, of course, that the proportionality is due to the fact that the gravitational field is

[22] See Maudlin (1989) and (1990).

[23] Butterfield (1989).

[24] Earman (1989), pp. 207–208.

[25] Maudlin (1990), p. 545.

[26] Ibid., p. 547.

[27] For a discussion of this argument, see Hoefer (2000).

[28] In his introduction to the Leibniz-Clare Correspondence, Alexander (1956), p. liv states two quotations of Einstein without any references, one in which Einstein says that the gravitational field "influences or even determines the metric laws of the space-time continuum," the other in which he maintains that the gravitational fields "define the metrical properties of the space measured." The first is from Einstein (1955), p. 62, whereas the second has not been possible to locate.

physically identical to the metric field.[29] Another possibility is to think of them as conceptually distinct but empirically identical. However, according to both Ruth Barcan and Saul Kripke, if such an identity proposition is true, it is necessarily true.

I believe that few would argue that inertia and gravitation are not conceptually distinct. But the *identification* of the intrinsic geometry of Space-time with the structure of the gravitational field cannot be an empirical discovery similar to the one that Hesperus and Phosphorus are the same. To see this we should first realize what it takes to be an empirical discovery. It means that observation brings together evidence that fulfills two different identifying descriptions. Ancient astronomers possessed different, empirically based criteria of being Hesperus and of being Phosphorus. But when it comes to identifying the metric of Space-time with the gravitational field, there are within GTR no such empirically based independent criteria of there being a definite metric structure apart from the gravitational field itself. We should also remember that the equivalence of the gravitational field and accelerated frames is merely local. This gives us problems with a global assignment of a unique metric structure founded on the gravitational field. Second, the effect of associating the gravitational field with the metric structure of Space-time itself is to physically narrow down the possible abstract space-time models that can be the model of the actual world. So the association is not an *empirical* identification but a metaphysical assumption that allows us to ground Space-time talk in physical reality.

Indeed, there is a sense in which inertia and gravitation are the same property that is only described in two different ways in different frames. The principle of equivalence ensures an account of the proportionality between the gravitational mass and the inertial mass, because it tells us that a system in free fall is an inertial system (locally). Therefore, the widespread understanding of GTR holds that the metric field (or it together with some related geometrical entity like connection) represents both the space-time geometry and the gravitational field. So when it is said that it has been *decided* by the physics community that it is meaning-

[29] However, Einstein saw GTR as a theory unifying inertia and gravity, not as a theory of geometrization of the gravitational field. Lehmkuhl (2014) strongly argues that Einstein, contrary to the folklore, emphatically believed that GRT "should not be interpreted as a 'geometrization' of gravity, especially if 'geometrization' was seen as a reduction of gravity/inertia to space-time geometry." (p.317)

ful to identify the gravitational field with the metric field, such a decision must be based on some assumption that is not part of any empirical discovery. Rather, the decision is based on a metaphysical assumption of co-existence, according to which it is physically impossible for the metric field to exist independently of the gravitational field.

This brings me to the second part of my argument. Maudlin considers the metric field as an essential part of space-time substantivalism. As we have just seen, the metric structure of Space-time is connected by necessity to the gravitational field where the notion of necessity is to be understood in a metaphysical sense and not merely in a physical sense.[30] Thus, Space-time is an entity whose existence cannot be separated from the existence of the gravitational field. So space-time points, and the metric structures we assign to these points, seem to be geometrical abstractions that do not exist independently of the gravitational field itself.

If this claim is correct, it is metaphysically impossible for a Space-time manifold to exist separate from gravitation. Therefore, I conclude that this four-dimensional representation of the world is a mathematical abstraction. Such an abstract conception of a Space-time manifold supplied with a metric and a topology is rich in structure and therefore it helps us to talk about the structure of the universe and a changing reality.

8.4 Space-Time Relationism

The proponent of the concreteness of Space-time is not limited to substantivalism. He could still argue that Space-time is a real entity but it reduces to spatio-temporal relations among the galaxies in the universe. But how can space-time points be concrete individuals without being a substance? The argument goes that space-time points are concrete because they owe their identity to concrete objects that occupy empirical space

[30] When Maudlin (1990) argues that "The substantivalist can regard the field equation as contingent truths, so that it is metaphysically possible for a particularly curved space-time to exist even if all of the matter in it were annihilated" (p. 551), he is talking about something else. Even if all matter is annihilated there still exists a so-called source free gravitational field that constitutes the metric field (see Norton 1985, pp. 243–244).

and time. Especially they owe their identity to continuants or rather physical events.

Relationism, however, does not fare any better than substantivalism. I shall not rehearse all the kinematical-dynamical arguments that have been put against it by Sklar, Friedman, Earman, and others. What is important for my purposes is that the relationist believes that Space-time does not exist over and above the concrete fields; he sees it merely as 'a structural quality of the field' and therefore claims that all talk about space-time points reduces to talks about a causal-equivalence class of events. By this founding maneuver the relationist finds space-time talk is as much concerned with concrete particulars as does the substantivalist. But the relationist's attempt to specify such an equivalent class of causally connected events suffers from the lack of a consistent criterion of identity that does not refer to space-time points.

Relationists claim that space-time points exist whenever events that occupy them exist. Thus, space-time points are concrete since they reduce to the concrete events in them. Space and Time are identical to the things and events that are commonly spoken of as 'in' Space-time. Events are then really constitutive parts of Space-time analogous to the way our arms and legs are not *in* our body, but parts of it, i.e., constitutive parts. However, I will show that this escape route provides no way out.

I suggest that we can have the sort of ontological reduction that the relationist needs to make her case only if a certain identity relation exists between the entity, which we want to reduce, and the parts to which we want to reduce it. The parts of a whole must not be exchangeable without the whole losing its identity. Thus, if a particular entity continues to be the "same" even if parts of it are replaced by different entities because the identity of such an entity is not dependent on the identity of the parts, then this entity is *not* reducible to the sum of its parts. For example, a human body does not consist of the mereological sum of its parts because the various organs may be transplanted by donor organs or artificial parts without the body losing its identity. In contrast, however, particulars like particular masses or quantities of stuff are numerically the same as the sum of their parts because they depend for their identity upon the identities of objects that are their own proper parts.

Although impossible to perform, it is possible to imagine that a planet, the Sun, or a galaxy could be replaced by another object of its kind and composition without Space-time changing its identity or geometrical description. Space-time would still have the same curvature everywhere and at every time. It would have the same metrical structure due to the same gravitational field, and it would still be a four-dimensional continuum. It seems at least that other material objects can substitute all individual objects, whereas the intrinsic properties of Space-time, which ground the identity condition of Space-time, stay the same all the way through the constant flux of matter and energy passing through them.

Indeed, there are less radical forms of relationism. One can argue: (1) that space-time points exist only in virtue of those continuants and events that occupy them, even though they are ontologically distinct from them, or (2) that space-time points exist only as possible places for continuants and events to exist. The metaphysical basis of the first claim is that an entity can be ontologically distinct from another entity only if they have independent identity conditions (as father and son). By making the identity of space-time points distinct from the identity of their occupants, but by claiming these points to be existentially dependent on those occupants, we only make a separation in thought, because their alleged distinct identity conditions have no empirical consequences. I conclude that this view collapses to the claim that space-time points are abstractions. In contrast, the second claim, however, presupposes that the possible places have some kind of existence independently of their occupants. This view therefore provides a way for a kind of substantivalism. Thus, none of the other forms of relationism do any better than the radical one at saving the presupposition that space-time points are concrete entities.

8.5 Space-Time as an Abstracted Concept

In my opinion, the traditional distinction between substantivalism and relationism is a false dichotomy: The distinction assumes that either (1) Space-time is an ontologically independent entity, because it has a structure that exits independently of physical things or events, or (2) it is reducible to the structural properties of things or events. But substantivalism and relationism are not contradictory terms; (1) implies that things or events are not *necessary* for the existence of Space and Time, whereas (2) implies that events or things are *sufficient* for the existence of Space and Time by presupposing that things and events are definable or identifiable without any reference to space and time. (2) expresses only *reductive relationism*, and one may deny (1) without being committed to (2). Things and events can be *necessary* conditions for the existence of Space and Time, even though physical space and time cannot non-circularly be defined in terms of things and events. I shall argue that the concept of physical space and time can be understood as abstracted from certain structural properties of the physical world as we experience it, and as such Space-time is an abstract conception by which we represent these things and events on a large scale.

Einstein's point-coincident solution to the hole argument was that two fields differing only by a diffeomorphism, like the metric field $g_{\mu\nu}$ and the dragged-along metric field $g'_{\mu\nu}$, are mathematically different, but in Einstein's words "entirely equivalent" and describe the same physical situation. Thus, it can be argued that Einstein's point-coincidence argument does not represent a view on the nature of space-time at all but merely an endorsement of the Leibniz equivalence. If this conclusion is correct, then the coincidences of space-time points do not just add up to a space-time structure. As Mauro Dorato emphasizes, "we certainly cannot reconstruct the global structure of a general relativistic space-time just by patching together such strictly local intersections of worldliness;" nevertheless, the local intersection of world lines provides the observational basis on which any theory of space-time can be built.[31] Geometry and pure theories of Space-time in general are logical or mathematical

[31] Dorato (2000). In a paper written together with Massimo Pauli, Dorato and Pauli (2007) argue for a theory named point structuralism which according to them is a combination of features from both substantivalism and relationism: "including elements common to the tradition of both substantivalism (spacetime has an autonomous existence independently of other bodies or matter fields) and relationism (the physical meaning of spacetime depends upon the relations between bodies or, in modern language, the specific reality of spacetime depends (also) upon the (matter) fields it contains)." (p. 147) They explain that their theory embodies entity realism as the metric field exists physically "as an extended entity together with its point-events." It is not just reducible to a mathematical structure. Furthermore, they claim that the space-time points exist, but their nature is relational. The effect is, if this is correct, that the metric field individuates the points of the manifold. See also Lusanna and Pauri (2006). Their view may be characterized as a form of non-reductive relationism; see note 32 and the discussion below.

abstractions useful for certain physical implementations, but I think that it is a serious mistake to hypostatize these abstractions. This view I call *non-reductive relationism*.

Non-reductive relationism takes the metric tensor $g_{\mu\nu}$ to represent a gravitational field rather than the Space-time structure itself.[32] Field theories seem to change the long-established debate between Newton and Leibniz. The non-reductive relationist does not have to fight the notion of empty space. There is no space where there are no fields, i.e., something physical. The attempt to maintain the classical perspective by defining the physical matter in terms of the matter-impulse-stress tensor $T_{\mu\nu}$ and then claiming that $T_{\mu\nu} = 0$ and $g_{\mu\nu} \neq 0$ represent empty space-time points is not convincing.[33] In general, $g_{\mu\nu}$ represents the gravitational energy, and the so-called vacuum solutions exist only in the real world as approximations where the source expressions are ignored. GTR is not a theory of matter, and the introduction of a theory of matter via quantum theory gives vacuum solutions different from zero.

Since reference to space-time points takes part in specifying the identity conditions of concrete particulars, i.e., world lines of galaxies, then Space-time itself cannot be a concrete particular. My suggestion is that Space-time is considered as an abstracted notion in the sense that it is conceptually dependent on fields and matter. Earman and others reach the substantivalist position by hypostatizing space-time points as objects, which are then thought of as the subject for predication of the properties of the fields that are defined at those points.[34] Here it seems as if Earman merely hypostatizes the diverse conceptual levels of differential geometry. We begin didactical-mathematically with a differential manifold, and

---

[32] Carlo Rovelli (1997), pp. 193–194, argues that Einstein's identification between gravitational field and geometry can be understood in opposite ways: (1) "the gravitational field is nothing but a local distortion of spacetime geometry" or (2) "*spacetime geometry is nothing but a manifestation of a particular physical field*, the gravitational field." He himself defends the second option, which I take to be an example of reductive relationism. The metric field is the manifestation of the gravitational field and as such "The metric/gravitational field has acquired most, if not all, the attributes that have characterized matter (as opposed to spacetime) from Descartes to Feynman." In contrast, the non-reductive relationist would say the actual geometry is an exemplification of infinitely many possible geometries and that physical space-time seems to gain individuality by being *instantiated* by the gravitational field.

[33] Friedman (1983), p. 223.

[34] See for instance Earman (1989), p. 155.

then we supply it with diverse affine, metric, and topological structures, and without any further argument it is taken for granted that this pure manifold exists ontologically independently of the structural features that characterize the world as we experience it. I hold that the problematic move in this line of reasoning is in the first place the very idea that we are allowed to reify space-time points as independent entities with their own criteria of identity.

Oliver Pooley takes issue with Earman and Norton's hole argument.[35] Following Gordon Belot and John Earman, he defines *sophisticated substantivalism* as any position that denies haecceitistic differences.[36] Such a position regards two diffeomorphic models as representations of the same possible world so they cannot be attacked by the hole argument. In contrast to Belot and Earman, Pooley holds the view that, as a sophisticated substantivalist, one can argue that space-time points are real substances, although their numerical distinctness is grounded by their position in a structure. He believes that such a modest structuralist position does not "go beyond an acceptance of the 'purely structural' properties of the entities in question," while at the same time maintaining that these objects cannot be *reduced* to the properties and relations themselves. I wonder, however, how space-time points, in terms of their mathematical structure, become physical space-time points. Pooley does not provide us with a single argument to show that the numerical distinctness of the mathematical objects of a manifold (points), whose identity depends on their positions in a mathematical structure, corresponds to the numerical distinctness of *real* physical space-time points.

Let me illustrate why I think to be problematic Pooley's suggestion that space-time points are real entities in spite of their purely structural properties. Take a series of identical billiard balls and add an ordering structure: one, two, three, four, five …, from the left, and then the identity qua 'number four from the left' is given in virtue of the entire structure, namely all the other billiard balls plus the given structure. But it is not a property of either particular ball that if we exchange it with the fifth ball, then the identity of the two balls switches, too. Each keeps its own

[35] Oliver Pooley (2006).

[36] Belot and Earman (2001), p. 228. See also Belot and Earman (1999).

identity before and after the switch, although the order itself is completely unaltered. Space-time points, however, are defined only in relational terms, meaning that they change identity whenever they change their place in the structure. Had they not changed identity and were they still individuated only by their place in the structure, then the order among themselves would not have stayed the same. Analogously, the identity of the number four is defined by its place in the entire sequence of numbers, and whatever whole number that may occupy the place between three and five would be identical with number four. Here both numbers and space-time points seem to be ontologically on par.

From the naturalistic perspective taken in this book, the characterization of space-time points as independently existing entities with their own identity conditions seems to be a problematic extrapolation of common sense ontology that assumes that in addition to their relations to other physical objects, all physical objects also have intrinsic properties that in part determine the kinds of relations into which they can enter. Space-time points lack intrinsic features, and without them there is no physical basis for differentiating one from another. The identity of space-time points as abstract objects is determined by the mathematical structure as a whole in the sense that we define and identify the constituents (points, etc.) within that entire structure. In such an abstract structure, the identity of any particular constituent is given by reference to all the other constituents plus a certain ordering relation among them. However, we cannot define and identify the entire structure in virtue of the structure itself. There is therefore a categorical difference between the constituents of the structure (points) and the structure as a whole. Their criteria of identity are not the same. We identify and define the constituents (points) within the structure, but such an individuation is not possible with respect to the structure itself. Nevertheless, from a naturalistic point of view, the claim that the identity of physical space-time points constitutes a primitive fact that does not require any further explanation is nothing more than an act of fiat.

Although I have sympathy for the spirit behind non-reductive relationism, I think that the entire discussion rests on the false assumption that GTR represents the global structure of Space-time as physical space-time in relation to the matter-energy distribution of the universe. In contrast, I believe that like every other scientific theory GTR does not represent the global structure of the universe but provides a vocabulary and a set of linguistic rules developed to establish particular models of gravitational systems. These models may be set up to represent the universe as a whole or any local system of gravitation in a local space-time region. In Chap. 4 I argued that structural realism in general holds an awkward position on the relationship between mathematically formulated models and the world, namely that there exists an isomorphic coherence between some mathematical structures, which exist independently of the world, and the real structure of the world as it exists independently of such mathematical structures. It does not suffice for the structural realist to point to the ontological commitments of the structures given to us by theories. The commitment to a certain structure is always internal to the mathematical framework according to Carnap's 'internal/external' distinction in semantics.

The structural realist needs to point to some external commitments that guarantee the existence of real physical counterparts. For instance, Dean Rickles argues:

> Structuralism has been 'creeping up' in physics for some time, and with the advent of gauge theories I think it is the obvious interpretative stance to adopt; when we turn to background independent theories like general relativity when even dynamics is pure gauge, then structuralism is almost forced upon us.[37]

However, as an evolutionary naturalist, I think that such a stance, which relies only on conceptual coherence and mathematical consistency, is completely misguided. Indeed, all kinds of questions about the existence of space-time structures *internal* to general relativity models are trivially answered with a "yes." But structural realists also claim to answer Carnap's *external* question. Here we want to know whether or not reality itself, not just our talk about it, contains Space-time structures, even though these structures do exist for all descriptive and reflective purposes. But as long as nobody has shown that the structure of mathematics has an ontologically different status than the structure of language, there are

[37] Rickles (2008), p. vi.

numerous unresolved questions that face those who would try to answer *external* questions about Space-time.

So far I have argued that space-time points and fields are ontologically distinct entities, because they belong to different kinds of existence, but I have also claimed that Space-time both ontologically and conceptually is posterior to changing and extended things. I now want to conclude that Space-time is nothing but a geometrically ordered *set* of space-time points to which we refer whenever we want to track down the development of fields and objects in a model. The geometrical conception of Space-time of GTR provides us with a mathematical language in terms of which we can talk about extension, relation, and their changes in relation to gravitation. We need the conceptual resources of Space-time to assist us in identifying and ordering things and events globally. We would be unable to represent particular things in a GTR model and to track them down from one moment to the next, unless we had the possibility of referring to their continuity through Space-time. Space-time gives us the conceptual tool to describe movement of the same material object through a variety of spatio-temporal locations and gives us the conceptual tool to talk about the persistence of numerically the same object possessing different properties from one spatial locus to another or from one temporal moment to another. Therefore, we may say that *particular* concrete objects are in physical space-time, meaning that their world-lines can be represented by a set of space-time points in a model. Thus, an enduring object is one that may undergo changes in physical space and time, while it continues to stay the same object throughout its changes. The question then arises: Is Space-time a mere linguistic or mathematical tool, a conceptual instrument for identifying concrete particulars, or does it have some sort of abstract existence?

I am inclined to hold that even *ordinary* space and time exist as abstractions, in the sense that their existence is ontologically, but not causally, dependent on the existence of concrete particulars organized by the mind. First, the mind grasps the concepts of spatial locus and temporal instant from the locations and relations between the objects of perception perceived as somewhere and from the order of their changes perceived as occurring at different intervals according to the internal causal order of perception. Next, this construction of local spatial position and temporal moment is eventually extended to all actual and possible events elsewhere in space and into the past and future in time. But such an extension requires a precise definition of simultaneity. Classical theories are based on such a definition, relativity theory on another. Hence, the concept of Space-time, as a mathematical generalization based on a physical definition of simultaneity in terms of light signals, is so constructed that it applies to the totality of all events and objects that exist in the universe from its beginning to its end.

So I advocate the view that Space-time consists of an ordered set of space-time points originally abstracted from perceptual awareness of extension, relations, and change. This construction is then applied to all changing things and therefore thought to represent physical space-time. Our conception of this set is acquired by empirical acquaintance with a limited number of members of the set, and their order is subsequently abstracted from their relations to all other possible members of the set. It does not follow from this that the term used for that abstracted *concept* refers to an abstract *object* over and above the entire collection of concrete members in the universe. However, being an ordered set Space-time exists as an constructed entity with its own internal identity conditions, and therefore Space-time is not reducible to a mere collection of events. As an abstracted concept, Space-time has no space-bounded or time-bounded properties; it is subject to only tenselessly true predication as far as its relational properties are concerned.

Thus, Space-time is not only a set, but also the ordered set of space-time points abstracted from concrete particulars and relations such that the ordered set of space-time points is intended to represent the set of concrete events in the universe. Any particular event may coexist with some other particular events, or precede or succeed some other particular events, and based on these facts we may assign a relation of simultaneousness as well as an order of being earlier or later to all these events. For this purpose Einstein chose light signals for a coherent but operationally unambiguous definition of simultaneity. In general, events causally (and perceptually) succeed each other and therefore belong to different subsets (hyperplanes) of coexisting events. The actual spatio-temporal order corresponds to causal relations between concrete things and events. By grounding the order of space-time points in the causal structure of some particular actual events, we are able to ascribe a unique and unambiguous order to all events in the universe. Every space-time point becomes ordered with respect to every other space-time point, and we may use this abstracted representation to identify uniquely any particular event in this ordered set. Indeed, Space-time is an abstraction that has a very privileged relation to physical reality. There are an infinite number of mathematical geometries, any one of which could represent the actual world, but which of them may be used to identify changes and movements in the universe depends on the distribution of fields and matter.

8.6 Are Space and Time Invented or Discovered?

But what is the foundation of our ordinary conception of space and time? If we characterize Space-time as a mathematical abstraction, which has been constructed out of our ordinary conception of space and time, and the notions of space and time stem from the sensuous presentation of concrete individuals, the ontological status of ordinary space and time is still unclear. Are space and time something that exist over and above those locations and relations we can perceive, or are they abstractions created by human intellectual faculties that have evolved through natural selection? Do we have a notion of empirical space and time because our mind has become biologically adapted to represent the external world as it really is, or do we have these notions because this is how the mind structures the world of experience based on how the external world is presented to our sensory faculties by millions of years of natural selection? The conception of space and time as something stretching beyond our immediate perceptual experience seems to have been both invented by human thought and discovered by reflection. Just as natural kinds are invented as the result of a cognitive capacity for similarity recognition, and later discovered by reflection, the external world is sensuously presented to us as consisting of similar spatial and temporal relations. As abstracted by reflection from memory and the sensuous presentations of the external world, these ordinary concepts have a content that goes beyond our experience, and even all possible experience, but nevertheless have the cognitive function of allowing us in our daily life to imagine motion and change and think about them as continuous processes. The same can be said about the Space-time conception of relativity theory that provides the physicist's currently accepted conceptual framework for representing the large-scale structure of the universe.

Often we associate objectively existing phenomena with things that can be empirically discovered (if they are humanly accessible), but those phenomena that exist only subjectively are associated with things that are a result of human invention and construction. But this is not always the case. Kant's notion of space and time as transcendental forms of intuition was a position that took space and time to be necessary a priori conditions of experience discovered by pure reason. Kant unequivocally rejected the empiricist view that they were abstracted from observed things. Contemporary space-time physics, interpreted realistically, also takes the structures of Space-time to be something that can be discovered. But, in contrast to Kant's a priori grounding, this time it is the result of an empirical discovery based on the experimental support of relativity theories.

Other things may be a result of invention and still exist objectively, that is, autonomously of the mind. A car and a computer are products of human invention but exist as physical things independently of the carmaker and the computer maker. But they exist as cars and computers only as long as we are around to assign a certain meaning to them in terms of their function. The same holds for languages, families, institutions, and states. But these do not exist as concrete particulars in space and time; they would not be here had it not been for human invention. The difference is that physical objects like a car and a computer ontologically depend on physical states of matter, whereas families, states, and the like ontologically depend on human beings and their communicative actions.

Thus, the question we have to address is whether or not space and time in the ordinary sense would have existed if human beings did not exist. Do they have a status as natural or artificial particulars or as nominal constructions? Space and time are the measures of extension, change, and motion. I have said that our notions of space and time, as abstractions, reflect our experience of the spatial and temporal relations among concrete particulars and therefore conceptually depend on the existence of concrete objects and physical events. The experience of spatial and temporal relations must have originally co-evolved with the perception and memory of the locations of their relata. Apparently, this indicates that space and time are discovered by us. But I also mentioned that our ordinary notion of space and time is abstracted from spatial and temporal perceptions whose content is mind-independent but not necessarily species-independent. Being species-dependent does not imply that spatial locations and temporal changes are brought into existence by us.

Another possibility is to say that the existence of ordinary space and time not only depends on the existence of concrete physical things, but also on the existence of human beings. To say that overall space and time depend on humans is not to say that they *only* depend on humans. We formed an ordinary concept of space and time, as well as a scientific concept of Space-time, through abstraction, and this process is indeed confined by the innate mechanism of our internal presentation of spatial and temporal relations among individuals in our environment. The concept is abstracted from a collection of common spatial and temporal patterns through which natural selection gave our ancestors the capacity to recognize among concrete existing individuals. What this concept refers to would be an abstract object if it had made sense to claim that such entities exist. As long as we do not hold that an abstract concept is fully determined by the existence of mind, but also by the existence of concrete relations from which the concept is abstracted, it does not have only a purely subjective origin. Our ordinary concept of space and time is concerned with the external world even though it is constructed by us based on perceptual and bodily information about relations among object in our environment. So it seems reasonable to maintain that space and time are neither discovered nor invented by us, since they would not have existed if it had not been for our physical environment and us. Space and time are *ontologically* species-dependent but *epistemologically* mind-independent, because we take them to supervene on spatial relations and temporal changes among concrete particulars, which form the empirical basis from which the concepts of them are abstracted.

I would say that the mental capacity for reflection allows us to construct space and time based on our experience of spatial and temporal relations. Kant was right insofar as it is by reasoning on the basis of these conceptual constructions that we come to know the geometrical properties of space and time. As intellectual abstractions ordinary space and time are not directly given in experience, although the process of mental abstracting, which produces our concepts, starts with what is given empirically. But Kant was certainly not right when he argued that space and time exist prior to and independently of experience as proven by his transcendental argumentation. We can experience spatial and temporal relations, and the way we get to the abstract notions of space and time is through a mental process starting with such experiences. We may think of ordinary space and time as names for constructions based on our perception of concrete particulars, and the reason why we have eventually formed a concept of them is to gain an ability to track down natural and artificial particulars beyond what is immediately given in experience. So our notions of space and time exist as abstract concepts, and their cognitive role is to provide us with an order for the assignment of truth values to descriptions that directly or indirectly refer to spatial and temporal relations among concrete particulars. So our concepts of ordinary space and time as such refer to neither transcendental nor real things-in-themselves.

Let me recapitulate the assertions of this section: we abstract the notions of what we may call "ordinary" space and time from our experience of the spatial and temporal relations between physical objects as presented in experience. So I hold that spatio-temporal relations are real features of the world quite apart from us, and our ordinary concepts of space and time are abstractions derived from our experience of these relations. So these "ordinary" notions refer to constructs, something we have created, i.e., they have no ontological status apart from human thinking about them. Normally we would say that makes them subjective rather than objective. But it seems also correct to say that they are not created by the individual subjective experiencer but by the species as a whole, so this gives "ordinary space and time" a reality apart from the referents of the concepts that exist in individual human minds. So "ordinary space and time" are neither objective (i.e., things in themselves) nor merely subjective (i.e., only subjective categories of thought) but have a reality dependent on the species as a whole, given its faculties as shaped by the forces of selection and survival. For this reason spatio-temporal knowledge is not subjective but has a universality and necessity for the species, even if it is not the universality and necessity of pure reason that Kant claimed. Thus, in a naturalistic account the effect of natural selection substitutes for Kant's grounding the forms of intuition in pure reason.

9

Causality and Counterfactuality

We commonly regard the concept of causality as essential for understanding the difference between the accidental and the non-accidental conjunction of events and the relation between the actual and the possible. Not only do we use causal notions whenever we try to grasp why something new or unexpected happens, but also our sensory experience of the most familiar events is conceptualized in terms of these notions. Our everyday world is a world in which we experience multifarious causal connections among the events we witness. Still, philosophers have raised doubts about the reality of causes as something over and above the regularities we experience among our sensations. Rightly or wrongly, David Hume was understood to claim that our experience of causation contains no necessity *in re*. All of our beliefs in causal modality stem from the mind's projection of necessity onto the events we experience. Only necessities *in dicto* exist. A born-again empiricist, like Bas van Fraassen, holds that "From an empiricist point of view, there are besides relations among actual matters of fact, only relations among words and ideas."[1] Modal and causal expressions, however, are not among those statements that are concerned with actual

[1] van Fraassen (1989), p. 213.

J. Faye, *Experience and Beyond*,
DOI 10.1007/978-3-319-31077-0_9

states of affairs, since they seem to be about relations among possibilities. Instead van Fraassen thinks that "a graphic, if somewhat inaccurate way to put this would be: causal and modal discourse describes features of our models, not features of the world."[2] In contrast to this hard-core empiricism, I shall argue that causal claims may refer to *real* relations, because causes are both actual and observable facts of the world. However, although modal features are part of our causal understanding, I do not believe that a causal claim entails that alleged modal entities such as possibilities and necessities have their own *real* existence.

Therefore, in this chapter I want to address the question of what the concept of causation includes beyond what originates in concrete sensory experience. It seems to be the case that biological evolution has installed cognitive schemata of causation in the consciousness of animals that enable them to coordinate bodily and sensory information, foreseeing what other animals can be expected to do and to behave according to the expectations. There seems to be a pretty obvious evolutionary explanation of why natural selection has favored animals that were conscious of causal connections in the world around them. However, such an awareness of causal connections does not imply that animals have modal concepts like necessity or that they can imagine counterfactual possibilities. That purely animal "causal schema" does not necessarily extend to a tool-using hominid who in some way grasps a connection between wielding the tool and its effects. I am tempted to say modal categories require some degree of self-consciousness (trans-animal) that allows us to ask: "Okay, *A* and *B* were temporally connected last time, will similar *A*-type events be connected to *B*-type events next time? And again and again…?" The innate schemata of causation constitute the capacity of animals to learn by induction how to perform in their environment. Nonetheless, the notion we have generated from reflecting on our own causal capacity seems to be more complex than what can be accounted for by an appeal to regularities.[3]

The world works in a regular way. And we are highly adapted to that fact. Of course two facts are true about human experience: (1) regularities

[2] Ibid., p. 214.

[3] Faye (2014), pp. 138–42. On page 136 I say "In Kantian terms the notion of causation is a posteriori for the species but a priori for the individual." By this I meant that the schemas of causation are innate and that an individual organism's actual notion of causation is its behavioral expression of the innateness of this capacity of forming causal beliefs.

happen over and over, and (2) occasionally genuine novelties arise. Thus, a well-adapted organism must be able to anticipate (Hume's habit) the regular reoccurrence of the regularities, but also must be able to somehow work in the genuine novelties into its worldview. To fail in either case would surely subject that organism to immanent peril. It seems to me that in a universe without regularities (one of those other possible worlds in a dark corner of the multiverse) evolution by natural selection would not work, since the mechanism or stratagem that was survival positive yesterday may not be tomorrow. To the extent that "our" world is indeed punctuated by cataclysmic changes, the delicate constructions of natural selection are wiped away and reset to begin all over again in a new world where old regularities may no longer obtain.

But philosophers often put more into the notion of causality than what can be accounted for by the regularity account of causation. The attractiveness of that account is its attempt to understand causation solely in terms of observational features of the empirical world. However, this account seems not to match our intuition that we experience many events as cases of singular causality that follow no repeated previously experienced cases of temporal succession. Since this experience is prior in many normal epistemic situations, it follows there is more to a reflective notion of causation than mere temporal succession. Apart from succession and dependency, the concept of causality is usually *defined* as containing a modal feature, which allows us to engage in counterfactual inferences about singular causal events and to claim that a particular cause is both sufficient and necessary for its effect in particular circumstances. However, we may directly observe what we regard as singular causes followed by singular effects, but the modal element of the alleged necessity of the connection is never something we can ever possibly observe. Rather, this element is something we add to our empirical perception of causal connections.[4] Thus, the present chapter suggests that the modal feature of necessity is

[4] Michael Ruse (1986) expresses the problem very nicely: "Certainly there are causal connections, and only a fool would ignore them. The point is that causes are not *things* (over and above the physical world), like powers or invisible fluids or such phenomena—though we have a tendency to think they are. Nor are there metaphysical hooks, or any such things, binding causes and effects. The world works in a regular way. It is in our biological interests to take note of this, and so as an adaptive response we tend to make something out of the regularities. But, as philosophers, we should not try to make more out of regularities than they are. Causes are projected into the world

a mental construction that was originally derived from our knowledge of certain structural features of similar events in different situations. It does not originate in what we actually observe, but is abstracted from what we have observed or may observe under different but pertinently similar circumstances. So the concept of causation has partly an empirical content and partly a constructed one. I shall propose that the empirical content is due to some causal schematum through which we perceive different events as causally connected, whereas all modal features that are part of our belief in that causal relation are constructed by the mind.

9.1 The Concept

Attempts to find causal mechanisms are already ubiquitous in our normal way of thinking, because causal understanding provides us with a well-ordered and structured comprehension of the world that provides many practical advantages in daily life. Expressing our recognition of causal connections is what causal discourse is all about. Causal language is used to communicate what we see as non-accidental connections in both everyday life and science. Fundamentally, there is no difference between causal claims made in science and in ordinary life. Scientific understanding differs from that of daily life only in that it is made on the basis of greater theoretical insights and is justified by more sophisticated means of experimentation than in the experiences of ordinary life.

Intuitively, we possess causal knowledge, because we experience the world as basically causal in nature. We observe causes: In the same manner as we experience things and events as having certain properties, we also experience these things and events as causally connected and as partaking in physical processes. Furthermore, we could not act in a world where we did not know—or did not have quite reliable expectations about—how things would behave. Our beliefs in causal connections help us to execute actions that will result in events we wish to bring about or to prevent events that we do not want to happen. Knowing causal

by us, through our epigenetic rules. The human who believes in real connections has the biological edge over the human who sees only contingency" (p. 174). I agree completely.

connections is also a precondition for doing experiments and for successfully making and operating technical devices like tools and instruments. Thus, as a matter of fact, most often causal beliefs represent genuine knowledge about the world. An obvious question is therefore how we obtain causal beliefs and what we come to believe about causation by obtaining it, another is what we mean by having 'causal understanding,' and finally there is the question of what constitutes causation as it is in the world itself, if there is such a thing. These three aspects of causation are more or less distinct.

The first one bears on the methodology of science, how, and the extent to which individual causal claims can be justified, the second is concerned with the nature of causal beliefs and the meaning of causal claims expressing those beliefs, and the third one deals with the ontological foundation of causal connections. I hold that in physics, for example, processes have a non-modal, or non-counterfactual, basis in the conservation of positive energy and its transmission in time and space.[5] But I believe that we should separate causation from physical processes.

In contrast to our understanding of physical processes, the meaning of causation is saturated with modal notions. The problem here is not that we do not understand causal statements—in fact we do, for we are able to use causal terms correctly without any hesitation. As competent speakers of a language we do not need a definition of causality to know what it is, any more than we need a definition to understand what human beings, animals, flowers, books, and cars are. We learn to use these terms correctly while learning our first language. Our linguistic practice shows that we are just as familiar with causation as we are with these other things. The interesting issue is indeed to get to understand what it is that we understand, and why we understand it as we do: what structure of the world, if any, lies behind the formation of causal concepts? What, in other words, is it about the world that has led evolution to produce *Homo sapiens* that experience the world through such causal concepts whose content supersedes what can have empirical support? The focus of our discussion will be the explication of what we mean by ascribing a causal

[5] Faye (1989, 1994).

nexus between events of certain kinds and how we can justify the modal interpretation of causal statements.

Philosophers have proposed a variety of different analyses that present a more elaborated account of our causal statements. The most common attempts are suggestions to analyze causality in terms of regularities, counterfactuals, or probabilities. All of these accounts have one thing in common, namely, the idea that the notion of causality is analyzable in terms of another notion (or notions), which is taken to be either more fundamental or better understood or directly experienced. In opposition to the reductionist accounts, I hold that causal claims cannot be fully defined in terms of any other concept. Indeed, all causal discourse is loaded with reference to regularities, counterfactuals, and probabilities, but these notions cannot by themselves replace the notion of causality.

Claims about causality imply claims about regularities, counterfactuals, and probabilities. Assume two individual events α and β are causally connected, then the phrase "α causes β" apparently implies the following claims: (1) events of type *B* follow events of type *A* given the actual circumstances *C*, (2) if α had not happened, then β would not have happened; (3) the probability of β is greater with α than without α. But at the same time causality gives rise to its own problems independently of any of these other related notions. Although none of these notions, individually or collectively, is able to explain causality, each of them may contribute to a philosophical elucidation of the concept. Nonetheless, I hold that the notion of causation is primitive in virtue of having evolved as part of our cognitive apparatus by natural selection. So what I propose is that although our understanding of causality is not completely analyzable in terms of regularity, counterfactuality, and probability, nevertheless it is so closely associated with them that it is impossible to grasp causality intellectually without taking its relation to these other notions into due consideration.

Thus, on the one hand, I agree with the non-reductionist conception that the truth conditions of singular causal statements are not completely translatable into the truth conditions of any other kind of statement. On the other hand, I subscribe to the antirealist view that the truth conditions of singular causal statements can be expressed in terms of what is epistemically accessible. Every part of the notion of causation, which goes beyond experience, and which therefore cannot be articulated in observational terms, does not correspond with anything objective in experience. I suggest that our understanding of causal claims is such that those parts of our concept of causation that include references to a necessary connection are due to abstractions from memory and perception. My contention is that it is only through our experience that we have acquired a grasp of the most general conditions under which a causal description of the world is possible.

9.2 Regularity

Causal thinking plays a major role in both science and everyday life, as we constantly acquire beliefs about the connections among events we experience in our daily surroundings. Scientists believe, for instance, that acid rain causes the death of trees in our forests and that PCP gases break down the ozone layer over the Antarctic. They also believe that the sun emits heat through the fusion of hydrogen into helium and that exposure to severe radioactive radiation produces cancer. In daily life we believe that the wind causes waves on the water and tree branches to sway and that exposure to water corrodes iron. But how can we be sure that these connections really exist; how do we prove that causal beliefs are not fictions but representations of something in the physical world? Why do we hold that we are justified in believing that the dependency between cause and effect is real, and not, as claimed by Hume, merely a projection of our subjective expectation onto our sense experiences?

Hume's account of causation treats it mainly as a regularity of similar events. As he says in his famous dictum: "We 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."[6] Singular causes reduce to generic causes, and generic causes reduce to a regularity of concomitant events. The idea of causal power according to which the cause brings about its effect seems to be part of our causal understanding. However, Hume thinks the basis of this idea is subjective, an expectation that we instinctively read into an objective feature of the phenomena.

[6] Hume (1748), sect. VII, part II, p. 76.

Perception does not support a notion of causal connection according to which the cause necessitates its effect in any sense over and above the notion of regularity where the effect merely follows the cause. What the latter does not capture is the psychological feeling of causal efficacy attending our perception, but that has no counterpart in sense impressions.

Following a Humean analysis, the traditional empiricist theory of causation is based on two independent conceptions: (1) to understand singular causal sentences means to realize that they can be derived from a general law; (2) the statement of a general law is merely a statement about regularity of the generic events in question. On the one hand, one may read "*X* causes *Y*" as stating a relationship between types, saying something like "*X* is that kind of event that forces *Y* to occur," "*X* produces *Y* in the proper circumstances," or "*X* is able to bring about *Y*." On the other hand, one can also read the sentence as expressing a relationship between tokens: "On this particular occasion an instance of *X* caused the instantiation of *Y*." Thus, empiricists hold that the second reading is a derivation from the first. Moreover, they usually regard the causal relationship between generic events as entirely constituted by their constant succession.

Sometimes we are able to observe that an event is caused by another, sometimes not. Indeed a person can be wrong about his perceptual judgment of causes as he can be wrong with respect to other kinds of perceptual judgments such as the predication of properties or the identification of things. The perception of causal facts is not different from perceiving other facts. Through our senses we simply acquire a belief that the term "cause" applies to a state of affairs in our sensory field. Thus, seeing what is going on does not depend on any insight in similar cases. A child needs to be burned only once to realize that the hot stove causes the burning pain. Quite often we see two particular instances of certain types of events succeeding one another for the first time; however, we still grasp them individually as causally connected. The idea of regularity between them as generic events is no part of the causal belief one acquires by merely seeing them as a singular instance of a causal connection. Rather, such a conception is inductively derived from the reiteration of individual but generically similar cases under certain recognizable circumstances. As a result we may eventually arrive at a general belief about causally connected events. In fact, a causal law can be tested only against particular instances of this law; therefore, singular causal connections cannot get their justification from the law that they are supposed to support.

Another problem is that in many cases a particular cause becomes an instance of a regularity or uniformity only because an unlimited number of exceptions, or *ceteris paribus* clauses, are included in the formulation of the causal generalization. There are, for instance, many children who get a stomachache from drinking milk, though millions of children drink milk every day without getting pains in their stomach. Thus, we cannot simply infer straight away from a number of individual observations to the generalization that all children drinking milk will get a stomachache. It has been discovered that some children are lactose intolerant, i.e., they are not able to decompose the lactose in milk; thus, drinking milk causes a stomachache. Therefore, one could formulate a more complex generalization, according to which drinking milk causes no pain except for lactose-intolerant people. But we can arrive at these complex and complete regularities only after further investigations and after having recognized the individual exceptions. Consequently, we do not need to know whether such regularities exist to be able to observe individual causes.

Apart from these shortcomings in the regularity view, the notion of singular causes does not entail the idea of a complex regularity for another reason: Individual causes can be associated with both the idea of positive and negative alternatives to what has actually happened. Different causes usually bring about different effects, but in any given situation the occurrence of an effect requires that its cause occurred. Whether this particular cause occurred or not makes an important difference in what became the effect: That particular cause was simply necessary for this particular event in the circumstances; it could not have been otherwise. As a consequence we want to be able to talk hypothetically about individual events. The effect would be absent if the cause had been absent; that is, the absence of the cause would be sufficient for the absence of the effect.

One regularity approach to causation is John Mackie's discussion of INUS conditions.[7] An INUS condition is an *i*nsufficient but *non-redundant* part of an *u*nnecessary but *s*ufficient condition. Mackie attributes this kind of analysis to John Stuart Mill. The latter pointed out

[7] Mackie (1972), 62 f.

two things: (1) an effect of a certain kind (*E*) is usually not preceded by merely one particular kind of events (*C*) but by a combination of several factors (*C* = *X*, *Y*, *Z*), that is, different types of events, states or situations that together cause the effect; (2) an effect can be preceded by more than one combination of factors. Several different sets of factors can generate the same effect. For instance, a fire may be caused by a short circuit, a bolt of lightning, or an arsonist. Thus, an INUS condition is an insufficient but non-redundant factor *X* that in conjunction with other similarly insufficient and non-redundant factors, *Y* and *Z*, form a sufficient cause *C* of *E*. Moreover, such a conjunction of INUS conditions may contain negative conditions, factors whose negation are conjuncts in a minimally sufficient condition.

What is called a cause will most often be the INUS condition that is most salient; the other INUS conditions are regarded as helping factors. That is so in everyday life, and that is so in science too. For instance, catalysts in chemical reactions may be regarded as the helping factors, while the salient factors are those that appear in the stoichiometric equation. But all INUS conditions can also be named INUS or partial causes. The full cause of *E* is then defined as the disjunction of all possible minimal sufficient conditions of *E* so that the full cause, in contrast to an INUS cause, becomes both necessary and sufficient for a certain type of effect.

The INUS condition account is a more sophisticated regularity theory of causation than Hume's old account. It opens up for the possibility of considering complex uniformities containing counteracting causes rather than simple uniformities. For this reason, regularity statements are often said to sustain counterfactuals, i.e., they allow one to infer them in connection with singular causal judgments. However, such generalizations are not entailed by these judgments. One consequence is that empiricists who consider the regularity theory as the proper account of every causal connection still have to face serious challenges. It seems they fail to offer a satisfactory explanation of the following two facts: First, the conception of singular causes is epistemically prior; that of generic causes comes afterward. We immediately have perceptual access to individual cases of causation, but we have no such direct awareness of general cases. We cannot get to the idea of a regularity of the same kinds of events under given circumstances without first having observed temporally and spatially contiguous instances of these events. Second, generally we believe that there is a fact about causally connected events further than mere succession. Causal connection is considered to be robust and not reducible to mere succession or a regular succession of similar events. Without a causal nexus linking the cause and the effect, the order in which the world has happened and is going to happen would be a matter of coincidence. Thus, we feel justified in talking about a single cause as necessitating its effect because we apparently understand that the effect must follow the cause in the given circumstances. In this context it seems also appropriate to talk about hypothetical alternatives: What would have happened in this particular case if the cause had not obtained? That is the reason why we feel justified in saying that a singular causal statement entails counterfactuals.

Even Hume seems in some way to admit that the causal connection is robust. For just after stating the above definition of causation he adds something that seems rather incompatible with the traditional regularity view of succession. It goes, "Or in other words where, if the first object had not been, the second never had existed."[8] How could Hume believe that this phrase is another way of expressing the content of his definition? Elsewhere, Hume seems also to turn against the notion of causation as a bare concatenation of events. In the *Treatise* Hume raises the question, "Shall we then rest contented with these two relations of contiguity and succession, as affording a complete idea of causation? By no means. An object may be contiguous and prior to another, without being considered as its cause. There is a NECESSARY CONNEXION to be taken into consideration; and that relation is of much greater importance, than any of the other two above-mentioned."[9] Perhaps Hume was not an empiricist but a skeptical realist as some recent scholars have suggested?

Simon Blackburn goes right to the heart of the issue, I think, when he argues against such an interpretation. Instead, he maintains that Hume wanted to distinguish between causation as a nexus holding between particular events and a straitjacket guaranteeing the continuation of the same pattern between these kinds of events again and again.[10] Whatever the

[8] Hume (1748), sec. VII, part II, p. 76.
[9] Hume (1740), p. 77.
[10] Blackburn (1993), pp. 94–107.

causal nexus is between two events on one occasion, the causal continuation of this matter of fact could in principle change, so that in other places or at other times the same connection between events of the same kinds might cease to exist (the problem of induction). The causal connection may be robust in the sense that the individual cause necessitates the individual effect, but this fact does not make it necessary that the events similar to the cause necessitate events similar to the effect. First and foremost, according to Blackburn, Hume debunked the idea that any inductive inference from one particular occasion to other similar occasions can be justified by appealing to straitjacketing powers or forces between those events. To generalize from particular experiences that the reign of such powers or forces can be extended to the future falls under exactly the same inductive limitations as the causal generalization itself. The knowledge of powers or forces cannot make certain that events would never occur otherwise than the way they have been observed to occur until now. Hume also argued that we can have no comprehension of the general idea of a relation over and above particular examples. Hence, he could not endorse a concept of law in which a natural necessity exists between universals.

In my opinion Blackburn's interpretation seems plausible. But, whatever Hume's view actually was, I agree that we have knowledge of particular causal connections independently of any grasp of laws of nature. But I disagree that we can have knowledge of particular causal situations without having an innate notion of causality. In my opinion this is the evolutionarily evolved cognitive "schematum" that we share with the higher animals, which also experience a world of causally connected events. Only we differ from them in that we have become self-conscious about this schematum and thus have a general concept of causality to reflect about. Dogs understand that certain events are causally connected; they just do not philosophize about it.

9.3 Modality

Unless we believe in a world where causal connections do hold between events, i.e., unless our cognitive apparatus includes an innate schematum of causation, according to which the effect is dependent on the cause, we will not experience these cases of succession of events as causally connected. And our evolutionary heritage explains why our cognitive apparatus operates this way. Moreover, the particular beliefs we hold about nature, i.e., our worldview, will determine the sorts of causal connections one experiences. The medieval mind "saw" Divine interventions everywhere in the course of nature; the Asian mind sees one's karma as causing all the events that happen to a person, etc. But the sense organs of these people act as do those of Western scientists who "see" none of these things. So we just cannot leave it up to sensory experience alone.

The basic notion of causality is not of our own making. In contrast to the empiricists' claim, I hold that our cognitive apparatus evolved with the capability of connecting events as cause and effect. Certainly experiencing a world in which events of certain types are causally connected did not just pop into evolutionary history with the advent of human self-consciousness. The notion of causality has been "constructed" but not by *us,* but rather by the process of natural selection over evolutionary history. When we see something acting as a cause of something else, it is not because we see directly one concrete event as necessary and sufficient for a successive concrete event. This is something we grasp, not from what we perceive directly, but from the manner in which the internal causal schematum functions by making some succeeding events dependent on other preceding events. When we do see one event as causally connected to another, it is because it fits the schematum of causality.

It might be useful to explicitly distinguish between (1) an evolutionary general schematum of causality that permits an organism to connect events causally and that evolved under the pressures of natural selection and so is shared with higher animals and (2) a reflective "concept" of causality that is possessed only by self-conscious *Homo sapiens* and has been constructed over cultural (or scientific) history. The former secures causal dependency among successive events in perception. The later form of construction is a cognitive process in which a new or a more complex concept is formed by considering concrete things and *adding* features that distinguish them from other concrete things. In this case the aim of construction is to make a distinction between accidental correlations and non-accidental correlations. Indeed, the addition of features beyond sensory experience should not be arbitrary. To avoid this we must look into what characterizes generically similar events we take to be causally connected in virtue of the application of the causal schematum. The starting point of the construction of a general concept of causality is to account for situations in which two actual events succeed each other non-accidentally. But we cannot move cognitively from the actual situation to *possible* situations without the use of abstraction. Before we can add features, we must subtract features. The method of abstraction is carried out by removing some of those features of the two actual events, which tie them to the particular actual circumstances of a single occurrence. As Nancy Cartwright's aptly describes abstraction with respect to causal laws, "It is not a matter of *changing* any particular features or properties, but rather of *subtracting*, not only the concrete circumstances but even the material in which the cause is embedded and all that follows from that."[11] However, getting to the concept of causation in the first instance I would prefer to say that we remove those features of the actual circumstances, which we recognize as being characteristic of only this particular context, by contrasting it with knowledge of the circumstances in which similar events succeeded each other and in which they did not. If these similar events appeared in a regular and predictive way, we said that the two actual events caused each other, whereas if these similar events did not appear in a regular way, we said that the two events did not cause each other.

Concrete events as such exist in space and time, and present events exist here and now; thus, similar events to the present ones exist at other spaces and times than here and now. So if causal connection is taken to be more than mere succession and dependency of two present events, we cannot directly observe in a singular particular case what the remaining features are. These modal features are not empirically accessible in any immediate way. This is so because they are added from what we know about similar events in corresponding situations. We make up those features in virtue of considering other relevant but different circumstances in which we could control similar events and intervene in their succession. The one of these similar events that is used to control the other is thought of as necessitating the other, and the one whose existence does not depend on the others is claimed to be necessary for the other. Thus, these structural features as

[11] Cartwright (1989), p. 187.

they are disclosed through control and manipulation with our environment become the modal features of causation, because even though the non-accidental patterns are observed for similar events, they are generalized to hold also in any actual case of succession.

In this manner our concept of causality becomes a cultural or philosophical construction made by self-reflection on our causal schematum. Our experience of a causal connection is not based on a projection of a psychological habit (as Hume proposed), but based on the work done by the schematum of causality (as Kant proposed in answer to Hume) we have as part of our evolution (while Kant would have said as part of our essence as rational beings). It is the modal part of our concept that helps us to understand what it means that two *possible* events may be causally connected. If we were not able to produce the complex idea that some possible events can be connected because the existence of one is necessary for the existence of the other, we would not be able to understand the difference between events that are possible by accident and events that are possible by causal necessity. The concept of causality is constructed by reflection to make a distinction between contingent dependency and contingent independency.

Granting that the reflective concept of causality contains modal elements, and granting that these do not correspond to ontologically real things in the actual case of causation, but epistemically represent the systematic occurrences of similar events in other empirically accessible situations, we need a way to talk about how the actual events would appear, if they were to be substituted for the similar events in these alternative situations. Instead of talking about generically *similar* events in other situations, we want to talk about the occurrence of exactly *same* events in other possible situations than the actual. Hence we invented counterfactual constructions.

First, the belief we have when seeing an event causing another event is not merely that these two events exist together or that they are somehow connected with one another. A central element of our causal belief is that one, the cause, is not only temporally but also *causally prior* to the other, its effect. Although cause and effect by definition are causally related, one of them is regarded as prior to the other. That is the cause. The effect comes about because of the cause, but not vice versa: The cause does not happen because of the effect.

Another belief is that the two events are connected because in actual circumstances the cause is *necessary* for its effect. For instance, warm weather causes the snow to melt, and in those situations the warmth can be regarded as causally necessary for the disappearance of the snow, in spite of the fact that other events can transform snow into water too. This feature of non-logical necessity is what is referred to by the use of counterfactuals. We express it by endorsing a sentence like “Had it not been for the thaw (and had all else been the same), the snow would not have melted.” It is obvious that the relationship between two accidentally co-occurring events cannot be described as obeying this kind of non-logical necessity.

The third belief commonly associated with the concept of causation is that causes are efficacious in the sense that they *necessitate* their effects. This is indeed another way of saying that causes are sufficient for their effects. Our concept of a cause is such that whenever we have a belief that one particular event is the cause of another we also believe that the causing event produces its effect by making its occurrence causally necessary. If something is a cause, nothing can stop it from being a cause by not letting its effect happen. For how can an individual event be a cause if it does not give rise to an effect? By definition a cause necessarily brings about an effect.

Ascribing truth values to counterfactuals goes beyond any possibility of actual confirmation. Quite evidently, counterfactual claims cannot be ascribed the value true or false on the basis of the observation of a non-existing situation, since the antecedent is supposed to be false. Understandably, this has troubled empiricists. But though we cannot directly confirm the truth value of a counterfactual statement by our experience of the actual situation, it does not mean that we are prevented from making any assessment of the truth value of it. For the necessity involved can be explicated in terms of possible worlds. Here I stand by the antirealist view concerning possible worlds as a family of models of our modal discourse. These worlds are nothing but fictitious constructions by which we can explicate our talk about necessity and possibility. But because possible worlds are not real, it does not imply that counterfactuals cannot be true or false, for given a certain possible world model of counterfactuals, our talks about the necessity or the possibility of the truth of counterfactuals can legitimately be justified by the instructions of the model in so far as the hypothetical claim of the counterfactuals can be actualized. Hence, a counterfactual claim about particular things or events should be understood as an abstract statement about how the entities mentioned will behave in other situations in virtue of their actual properties. In other words, the truth-makers or ontological ground for imaginable counterfactuals should always be found in some factual, or categorical, properties of the things or of the events in question.

The claim just mentioned is not without problems. For how can counterfactuals whose antecedent and consequent are actually false be true because of some factual properties, unless we allow some reification of possibilities (and therefore possible worlds)? We say that counterfactuals are true just in case the hypothetical content is imaginable but this is not the same as stating that it is realized. Consequently, there cannot be categorical facts of the matter that make them true or false. Although this conclusion seems straightforward, I do not think it is inevitable. My reason for doubting the inevitability of this conclusion is that counterfactuals function very much like indexicals in the sense that they contain an implicit reference to both a certain moment of time t and certain particulars $x_1, x_2, \ldots, x_n$. When someone utters, contrary to the fact, that "Had the weeds in my garden been sprayed with herbicides, then they would have died," she has a certain period of time and certain particular weeds in mind. It is with respect to these moments and these individuals that the facts of the matter expressed in the antecedent and the consequent are not realized. But nothing dictates that the actual grounds for ascribing a truth value to this counterfactual should be confined to the intended interval or the intended particulars. Here we must distinguish between two views concerning the grounds, corresponding to strong and weak constructivism respectively.

The first position holds that the grounds have to cite only those properties that are actualized at a certain time t equal to the intended time and satisfied by certain particulars equal to the intended particulars. As a result counterfactuals about the empirical state of affairs cannot literally be true or false. The second position holds that the grounds only have to refer to what is actualized at any time different from t, the implicit time in question, and are satisfied by any particulars different from $x_1, x_2, \ldots, x_n$, the implicit particulars in question, but particulars of the same sort. So this view claims that counterfactuals concerning empirical matters are true or false if their hypothetical content can somehow be actualized at any time other than *t* and that it is in fact realized at a time t^*. The two kinds of counterfactuals assumed to be about empirical state of affairs but that do not have a factual content—and therefore no proper truth value—are those counterfactuals whose hypothetical content is not realizable, since they contain an explicit reference to a certain time or particular, or which, for whatever reason, was never realized. In either case, no fact exists that can make such counterfactuals either true or false.

In his study of counterfactuals in terms of possible worlds, David Lewis argued that a counterfactual is true if, and only if, every world where the antecedent as well as the consequent holds is closer to the actual world than any world where only the antecedent holds.[12] However, philosophers have questioned Lewis' account because this specification of the truth conditions of counterfactuals suffers from the weakness that we have no precise way of defining which world is closer or similar to which. Apparently, we cannot even characterize the similarity relation in terms of facts about these worlds, say, which laws of nature belong to the closest worlds, since Lewis is ready to sacrifice what we believe to be the laws of nature to save the asymmetry in the evaluation of counterfactuals. I believe, nevertheless, that it is possible to give a coherent argument in which some of the empirically accessible situations (worlds) are closer to the actual situation (world) than others.[13] The cost is that we cannot preserve any asymmetry in the relation between the content of the antecedent and that of the consequent except what rests on temporal succession.

Let us distinguish between the *actual* circumstances and what could be called the *causally relevant* circumstances. The actual circumstances are, for instance, those that are present whenever an event causes another, whereas the causally relevant circumstances are those that are present each time similar events co-occur (i.e., those conditions that are present in every possible world in which events of the same sorts as the actual ones co-occur). So the actual circumstances include all causally relevant circumstances, but the converse is not the case: the causally relevant circumstances do not include every actual circumstance. In addition to the causally relevant conditions, the actual circumstances consist of conditions that are

[12] Lewis (1973), pp. 16–18.

[13] Faye (1989), pp. 65–74

causally irrelevant, those being events that do not repeatedly occur every time events similar to the actual events co-occur.[14] Hence, any situation (world), consisting of circumstances that are taken to be quite similar to the actual circumstances, is closer to the actual situation than every situation that consists of the causally relevant circumstances, but not of every circumstance equal to the actual circumstances. Analogously, every situation consisting of the causally relevant circumstances is closer to the actual situation than any that does not contain all of them.

Counterfactual statements we make about individual events involved in a causal connection are warranted only because we always recognize similar events as instances of generic events and therefore name them with the same kind of terms. The separation of the world into particular events and things presupposes a distinction between tokens and types, as well as a knowledge allowing us to apply type names to individuals in the correct circumstances. In my opinion, ascribing a truth value to a counterfactual concerning empirical state of affairs, like "If the gas had not been lit, the water would not have boiled," can be justified by simple induction from the observation of similar situations in which the circumstances are generically equivalent to the actual circumstances, except for events of the same sorts as the cause and the effect. We are observing such a situation whenever we realize that a stove is working, a kettle filled with water placed on the stove, and oxygen present without seeing an event similar to the cause (lighting of the gas), and an event similar to the effect (boiling water). Similarly, I argue that in those precise circumstances, seeing no event of the same type as the effect amounts to seeing no event of the same type as the cause. Thus, on the basis of the same observations and simple induction we may also evaluate a statement like "If the water did not boil, then the gas would not have been lit" as truthfully stated.

So we can conclude that experience yields support for a counterfactual symmetry and not the asymmetry we suspect the causal connection to have.[15] This implies, if it is true, that counterfactuals cannot explicate the entire meaning of causal statements. However, causal claims support counterfactual claims. Since empirical observation would assign a truth value to

[14] Ibid., pp. 160–163.

[15] Faye (1994), pp. 143–147.

both a counterfactual and its converse, one could in those cases infer the truth of the one from the truth of the other. The idea that these events are *causally connected* seems nonetheless adequately caught in that language.

Summing up, I have argued that our basic notion of causality is a "construction" by nature in the sense that cognitive selection and adaptation rest at the heart of the origin of this notion. The human cognitive apparatus operates with something like an innate schematum of causal connection in which some adjacent and successive events are seen to be dependent on others. As an evolutionary naturalist, I agree with those empiricists who believe that modal talk does not refer to anything in perception. Not being parts of the causal schematum, 'necessity' and 'sufficiency' do not belong to the natural mechanisms of cognition. As conceptual inventions, the concepts of necessity and sufficiency belong to the epistemology of reflection and not ontology. However, hard-core empiricists are "actualists," by which I mean that they allow only the present situation to be capable of generating a causal belief and conferring truth value on that belief. But if we are pragmatists of a more evolutionary sort, we do not have to restrict our causal claims to what we actually observe, but to what we have observed and what we may observe. Here modal concepts like 'necessity' and 'possibility' help us to grasp what we are now actually observing in the light of what we have observed in the past and to reason about what we can expect to see in the future.

It is obvious that science is predicated on the presupposition of what is usually called "the principle of the uniformity of nature." By nature we instinctively *presuppose* that the future will resemble the past, and this tendency also forms the foundation of the modern explanation of biological evolution. But the only "justification" we can offer is neither empirical nor a priori, but pragmatic: the success of the science (or the scientist) that presupposes it is the justification for that very presupposition. If that vaunted success of science comes to a halt and scientific predictions begin to breakdown on all sides, then we would have grounds to doubt our presupposition of uniformity. But insofar as that has not yet happened, we remain justified. While for empiricists it is anathema, for pragmatists the notion of innate mental structures is detached from bedrock reason and reattached to pragmatic, i.e., empirical, success. And evolutionary naturalists have a perfect and obvious Darwinian account of why our cognitive apparatus so operates.

10

Human Evolution and Mathematical Physics

The association of the aim of science with the discovery of mechanical laws formulated in mathematical terms is part of the historical legacy of the Renaissance. Such laws of nature were taken as the basis of science's explanation of why particular things, or phenomena, are what they are, or occur as they occur. Thus it must be an important aim of science to reveal as many laws as possible. Whenever we observe things behaving in a regular way, we may think this is because laws of nature exist as physical structures independently of things they link together, or we may consider this is because laws of nature supervene on their causal powers and properties.

A familiar opinion, shared by metaphysical realists with a Platonic orientation, is that laws are ontologically prior to things and events, and the relations and properties of these things would not be what they are if there were no laws. Laws of nature link things together, forcing things and events to obey them. Thus, if it is a law of nature that *A*, where *A* is either a certain state of affairs is the case or a certain events that happens, then it is not only the case that *A*, but also it is physically necessarily the case that *A*. In general laws of nature tolerate no exception. Since it is impossible for anything to deviate from their rule, science can use laws to explain why sorts of things in connection with other sorts of

J. Faye, *Experience and Beyond*,
DOI 10.1007/978-3-319-31077-0_10

things happen in the way they do. From an epistemological perspective we might get to know regularities first and then work our way toward laws. But from an ontological perspective the opposite is the case: Laws come first, particular regularities second; they are the consequences of the interplay of laws.

Another branch of metaphysical realism regards laws as supervening on the modal properties of particulars. Such properties could be in the form of deterministic dispositions or objective propensities that can be attributed to individual physical objects. Although this latter version of ontological realism may keep laws of nature distinct from their mathematical formulation, the former version, which considers laws of nature as ontologically primitive, takes them to be intrinsically mathematical. Indeed, an evolutionary naturalist opposes both of these versions. In the preceding chapter we assessed the foundation of the latter version, but in this chapter we shall focus on the Platonic version.

In its Platonic form metaphysical realism assumes that laws of nature are nothing but mathematical structures described by mathematically formulated statements. Today, advanced physicists search for algebraic features of various group theories to discover something physically about the world. A common argument in favor of mathematical Platonism is that mathematical objects and structures are indispensable as truthmakers for these kinds of theories to be successful.

Although this image of mathematically expressed laws of nature is prevalent among some physicists and philosophers, there are good reasons to resist it. Rejecting the existence of abstract objects implies the untenability of a strong realist notion of laws, in which laws of nature are taken to be relations between mathematical particulars. Therefore, our central challenge is whether the widespread success in science of explanations employing the notion of mathematical laws implies that something exists over and above regularities among concrete particulars or is due to a mental abstraction based on our reification of observational regularities among events. In other words, do the apparent referents of mathematical laws, assuming they are true, exist in their own right as ontologically independent entities, or are these laws merely a way of symbolizing relationships among classes of empirical particulars? Those who hold that laws of nature are ontologically different from the empirical manifestations of them usually claim that they relate to their instances with physical necessity or with some specified degree of probability. How can we justify the mathematical nature of physical necessity and probability in case mathematical laws are not identical with the actual manifestation of the laws? An evolutionary naturalist would presumably say what happened was something like this: organisms early on developed capacities for identifying kinds or classes of empirical stimuli; they could differentiate organisms-with-which-I-can-mate from ones-that-are-not-of-my-kind; they could distinguish food from non-food. The next step is that some organisms would develop the memory capacity to permit the ability to recognize patterns in the succession of classes of empirical stimuli; this capacity would be selected as survival positive. As organisms became conscious to higher and higher degrees they came to have expectations about the patterns of empirical stimuli. Every new sunrise was not a cause for astonishment, etc., etc. Since all you need to tell this story is empirical stimuli received by organisms of greater and greater degrees of consciousness—all the way up to self-conscious *homo*, evolutionary naturalism need postulate nothing more to account for the success of scientists' use of mathematical laws. No Platonic entities are necessary.

10.1 Mathematics and Representational Knowledge

For hundreds of years epistemological questions have been shaped by the way in which Descartes anticipated them and gave his answers to them. Looking back at the Cartesian theory of knowledge, we can see that it was occupied with three major questions: (1) What is knowledge? (2) How is knowledge possible? (3) What should we do in order to obtain knowledge? The background for these questions was his conviction that mental states, associated with sensory qualities, acted as unreliable internal representations of the external world. So how could we be certain that our senses and imagination give us a picture that corresponds to anything in the real world?

Regarding the first question, Descartes responded that a belief counts as knowledge if it is a fundamental belief or can be logically derived from a fundamental belief. A belief is fundamental, according to Descartes, if it is deduced directly from those ideas that appear clear and distinct to the faculty of "rational intuition." There are certain beliefs (or other mental states), which in virtue of their nature alone—their contents—have such a character that they are epistemologically basic. They are trustworthy because their truth is self-evident.

Descartes understood the second question as one concerning how we could avoid skepticism. For him an answer to this question meant that among those ideas that are clear and distinct he could find one that might help him to block skepticism. He found this idea in the concept of an infinitely perfect God who would not deceive him, because part of God's perfection is that he necessarily exists, and due to his perfection we are justified in believing in the certainty of judgments based on clear and distinct ideas such as those discovered in mathematics. Otherwise God would be guilty of having given us defective faculties, but that is incompatible with His infinite perfection. So even though our empirical beliefs about the external world may be in error, we could not be skeptical about the knowledge we have of mathematical propositions. Thus, due to the insight of mathematical intuition we have an essential insight in nature and with it the certainty of the mathematical description of an external world that came to be known as the mechanical world view.

Finally, he answered the third question by pointing to the fact that beliefs that were deductively inferred from this basic epistemic foundation would have the same epistemic status as the foundation itself. So when we put our clear and distinct ideas together we can infer knowledge about the world, which is guaranteed to be true. The Cartesian view does not hold that our experience of the world represents the world as it is in-itself; it is the Cartesian mechanical-geometrical model of reality drawn from the mind's stock of innate ideas that "represents" the world as it really is. The Cartesians deny a "naïve" form of empiricist representationalism, but accept a "sophisticated" form of rationalist mathematical representationalism. Hence, Descartes' representational theory of knowledge is very far from a naturalistic understanding of human knowledge as delineated above.

It seems quite obvious that very few of our empirically acquired beliefs can meet Descartes's standards of knowledge. Scientific beliefs are inductively established or sustained by inductively collected evidence.

Nevertheless, the Cartesian physicists believe in a strong form of mathematical foundationalism. Instead of seeing mathematics as a language of calculation and communication, they give it the role of revealing the truth about the essential nature of physical objects in the world. They consider mathematical knowledge to be based on the clear and distinct ideas that Descartes demanded for human knowledge, so they can infer from the mathematical representation of nature to the essential nature of the physical world. (For Descartes, of course, there is also another world of thinking beings that is not described mathematically.) The most provocative view among neo-Cartesians maintains that the physical world is mathematical by nature. The real world behind its physical appearance is nothing but an abstract mathematical structure. The assumption underlying such a claim seems to be that mathematical sentences are necessarily true, which means that mathematical knowledge is self-evidently true. It is logically impossible (i.e., inconsistent) to be skeptical about mathematical truth. Hence, modern Cartesian physicists do not need God; their mathematical intuitions guarantee that an omniscient mathematician could calculate how reality really is. If we know these mathematically formulated laws of nature, we can in principle deduce all empirical knowledge from a small stock of basic mathematical concepts. Whatever can be consistently described by a mathematical theory is certain to be true. Perhaps not in the actual world, but then in another possible world!

The indispensable presupposition lurks behind the representational view of mathematical theories. It maintains that we are ontologically committed to those and only those entities that are indispensable for our best scientific theories, and since many of our best scientific theories are formulated by means of mathematics, the reality of mathematical entities must be posited in order to explain the success of our best scientific theories. Therefore, we must have ontological commitments to mathematical entities. In the extended Cartesian version the argument runs like this: in physics, as an example, group theory is essential because various groups describe the symmetries that the laws of physics seem to obey. Noether's theorem states that any continuous symmetry of a physical system corresponds to a conservation law of the system. In general, physicists are interested in group representations, especially of Lie groups, since these representations often point to the foundation of possible physical theories.

Examples of the use of groups in physics include the Standard Model, gauge theory, the Lorentz group, and the Poincaré group. Thus, if a physical theory can be derived from an algebraic group alone, it seems to guarantee that the appeal to the mathematical entities of the group may explain the truth of the fundamental physical laws.

Usually nominalists have countered the indispensability argument in two ways. Either they have argued that we can get rid of mathematical expressions by paraphrasing them in terms of sentences without any reference to numbers. Or they have argued that numbers by themselves do not have any explanatory force. An excellent approach in using the first strategy has been developed by Hartry Field who proved that Newton's theory of gravitation can be formulated without numbers.[1] Although Field's paraphrasing was successful in the case of Newton's laws, it has by no means been demonstrated that it would be possible to do the same for all physical theories. Other objections point to the fact that numbers such as negative, irrational, and imaginary numbers as well as transfinite numbers cannot be paraphrased. The second strategy is to maintain that the mathematical components of our best theories do not explain anything. They merely serve as symbolic devices by which we can designate particular features of the world. Joseph Melia takes numbers to be a form of "indexing": If we say that the distance between *a* and *b* is 7/11 meters, we do not claim that the three place relation between *a* and *b* and the number do any explanatory work. "Rather, the various numbers are used merely to index different distance relations."[2] According to Melia, numbers function as the scaffolding upon which the bridge must be built, but when the bridge is finished the scaffolding can be removed. In another place he puts it, "When we come to explain [a physical fact] *F*, our best theory may offer as an explanation '*F* occurs because *P* is √2 meters long.' But we all recognize that, though the number √2 is cited in our explanation, it is the *length* of *P* that is responsible for *F*, not the fact that the length is picked out by a real number."[3] Our language has evolved in such a way that we use numbers to give a physical description of length, but we need have no ontological commitments to the facts of how we pick out length.

[1] Field (1980), pp. 61–91.

[2] Melia (2000), p. 473.

[3] Melia (2002), p. 76

These attempts to disconfirm the indispensability argument have produced an interesting response from Alan Baker. He uses the example of the genus of cicadas, called *Magicicada*, which spends most of its life underground as nymphs until at 13 or 17 years they all appear synchronously.[4] One possible explanation is that the emergence period of larger prime numbers was a predator avoidance strategy adopted to eliminate the possibility of potential predators receiving the same periodic population strategy. In this case, Baker claims, an explanation involving a 13- or 17-year period cannot avoid pointing to the mathematical fact that these are prime numbers. However, does this fact imply that ontologically there are numbers and that these are abstract entities? Or does it imply that there are empirically distinguishable entities (whole intervals of 13 and 17 years) that can be numbered and as such they are not always empirically divisible into entities (a series of intervals of one year) having a lower number other than the one that is picked as the proper entity to be numbered? I shall argue for the latter. The "number of objects" is quite arbitrary depending on how our ontology chops up reality.

10.2 Mathematics—The Language of Quantities

Metaphysical realists may seek a further argument to defend their view that the laws of nature have an ontological status over and above the manifestation of empirical regularities. Some may argue that because physical laws often make indispensable use of mathematics, such laws need to be something other than conventional constructions or empirical generalizations. In order for the natural sciences to apply mathematics to the physical world, the mathematics must have truth conditions that can be satisfied. Hence, abstract objects such as numbers must exist, because they are the truth-makers of mathematical propositions. In itself mathematics is concerned with abstract entities that are the truth-makers of mathematical sentences. Therefore, realists argue that whenever mathematics can be used to articulate laws of nature, those descriptions must refer to mathematical structures that instantiate mathematical universals.

[4] Baker (2005), pp. 229 ff.

Higher animals can count, nonetheless. Some birds and mammals can count accurately up to a very limited number.[5] They have evolved this capacity as part of their visual object tracking; that is their way of determining the number of separated objects. For an evolutionary naturalist this is the sensory basis for the use of numbers. The cardinality of a set is the final non-verbal tag that an animal can put on the objects counted. Moreover, animals have different mental representations of number 4 or 5. Some can even calculate small numbers in the form of addition, subtraction, multiplication, and division. Although animals have a primitive number system, and primitive abilities to calculate, they are nothing compared to human beings. Not only can we count to whatever large number we want, but mathematicians have also developed sophisticated method for handling numbers. But in this case, the realists argue, the nominalist cannot explain the use of numbers by just mapping numbers with sets of concrete objects, because there are transfinite cardinals of an infinite set, and the actual number of objects in the universe is apparently finite. If nothing else is an obvious candidate for being a real abstract object, numbers seem to be what we are looking for. The evolutionary naturalist cannot appeal to traditional trope nominalism that attempts to reduce universals to abstract particulars. Nonetheless, he has some suggestions to offer. Our counting skills and ability to calculate stem from our capacity for tracking and separating visual objects. This forms the natural foundation of mathematics; apart from that, mathematical entities are structures of constructions. Just as a novel may incorporate a few facts of life, the main part can still be a result of linguistic and narrative conventions and individual thinking.

Against any attempt to naturalize mathematics by considering it as having the same status as natural languages, namely as a vehicle of communication, an opponent of naturalism could say that even a natural language contains a lot of denoting terms such as names, nouns, and noun phrases. However, number terms used in ordinary language seem to play different roles depending in the grammatical form of different sentences. Consider an "innocent" sentence like:

[5] See Shettlewoth (2010), Chap. 10, and the many references therein.

(1) Jupiter has four (Galilean) moons.

In this “Jupiter” and “moon” denote physical entities, whereas “four” is an adjective and acts as a quantifier or determiner like “some,” “all,” “many,” “few,” etc. But if we consider a semantically equivalent but metaphysically loaded alternative sentence such as

(2) The number of Galilean moons of Jupiter is four,

then “four” no longer stands in the position of being an adjective. Apparently, (2) states that the definite description “The number of Galilean moons of Jupiter” is identical with “four,” and since singular terms such as definite descriptions are referring, the term “four,” having the syntactic position of a singular term, seems to refer to some object too. In (2) the singular term “four” seems to denote a particular number separated from physically being a Galilean moon.

Naturalists argue that the solution to this puzzle is not to claim that numbers are real after all. As Thomas Hofweber argues, the difference between (1) and (2) is due to a difference in the structural focus, and the difference in the structural focus depends on the context of communication in which these sentences are used. If you are in a situation where you want to tell me that Galileo already discovered moons around Jupiter, a statement like (1) seems to be an appropriate way to say that Galileo attributed some moons to Jupiter. But if instead you had stated that Jupiter has three Galilean moons, a more knowledgeable person may then correct you by uttering (2), perhaps with an addition “not three” at the end of the sentence. This person would now put our focus on the number rather than the moon or its discoverer. Hofweber concludes his analysis of how number words function in ordinary language: “‘four’ is not a denoting expression in either the innocent statement or in the loaded counterpart. That structural focus arises from extraction and movement shows that it is the same word, ‘four’, with the same semantic function…”[6] And he continues: (1) and (2) “have the same referring or denotational terms. There is no new referring terms coming out of

[6] Hofweber (2007), p. 23.

nowhere in [2]. 'Four' is merely moved into a special syntactic position in order to achieve structural focus."[7] Hofweber also shows how a similar analysis can be extended to properties and propositions.

If we accept Hofweber's account of the linguistic significance of number terms, then how can we explain some non-linguistic beings' ability to count and recognize small numbers? For nominalists counting does not involve a reference to abstract numbers but is a form of hashtagging, labeling the objects of observation with a mental tag, such as this is #firstLion, #secondLion, #thirdLion, #fourthLion, etc. The mental tags label individual lions, indicating nothing but that this particular lion has been labeled number one, two, three, etc. Natural selection, the evolutionary naturalist claims, selected these abilities because they were inherited and they gave a reproductive advantage to those organisms that possessed these abilities. The last number tag assigned to an object in a counting process defines the cardinality of the counted set, and whenever two counted sets have the same cardinality, it means the last assigned number tag in the counting process is "the same" in both cases. Like hashtagging on the social media, where the hashtag signals "open me," the outcome of number tagging guides the action of the tagger between engaging (a fight) and avoiding (a flight), between chasing and stop chasing, etc. As time passed and *Homo sapiens* evolved, its members eventually developed words to stand for mental number tagging together with a natural language. It is reasonable to assume that number words became part of natural language long before the development of a rational mathematics with its own rules and words.

It comes as no surprise that like ordinary language mathematics contains structure. Mathematics involves much more than simple counting, just as natural languages involve much more than labeling objects with words. Both ordinary language and mathematics contain rules for combining words together and numbers together. This is what gives us linguistic structure as well as mathematical structure. Apart from numbers used in counting, mathematics consists of developed methods for handling numbers, that is, various rules for making inference by which a set of input numbers is transformed to another set of output numbers

[7] Ibid., p. 24.

in terms of different mathematical functions. Thus, different functions defined on the continuum of real numbers or imaginary numbers give us distinct mathematical structures. From a naturalist perspective I would conclude that *we* use mathematics to express our (true) beliefs about the physical world, just as we can use ordinary language to express other (true) beliefs about the physical world. But mathematics in itself, just like the ordinary language in itself, does not represent the structure of the world.

10.3 Possible Worlds, Many Worlds and Multiverses

The metaphysical realist may still have an argument hidden up her sleeve. Laws of nature themselves exist not only actually, but law statements express a nomic necessary relation or a nomic probabilistic relation. Since the metaphysical realists think that this modal discourse is true, she is forced to provide a metaphysical account of the nature of its truth-makers. Modal notions can be given a realist interpretation in terms of possible worlds. A metaphysical realist will think of a possible world as a real world that is spatial-temporal isolated from any other possible world and therefore also causally separated from each other. The notion of probability can then be construed such that any probability is a quantitative measure of a possibility. This means, of course, that whenever physicists talk about probabilities, they are in fact referring to what is going to happen in some other real worlds apart from the actual one.

Hence, some philosophers and physicists are alike: they believe in the existence of other worlds is just as real as the actual world. Some philosophers argue that these other possible worlds must be real in order to make sense to our talk about necessity and possibilities; some physicists maintain that, due to the present of nomic probabilities, the measurement problem in quantum mechanics can be understood only if many worlds are introduced as real, and again some "daffy" cosmologists believe that a multitude of universes is the way to make sense of the so-called anthropic principle. Those philosophers and physicists who take such a stand believe that they make a substantial existence claim. They are realists about these worlds. So my job as an evolutionary naturalist is to scrutinize the arguments for this form of modal realism.

Notably David Lewis has defended possible world realism. He does not consider 'necessity' and 'possibility' as primitive. In a sense his view is reductive because he paraphrases talk of necessity and possibilities in terms of talk about possible worlds. Therefore, he must find a way to express what is possible and what is not, and he finds this in the principle of recombination: "according to which patching together part of different possible worlds yields another possible world. Roughly speaking, the principle is that anything can coexist with anything else, at least provided they occupy distinct spatiotemporal positions."[8] But how does he handle physical necessity and nomic necessity?

Philosophers have for a long time speculated about the reality of other worlds, but one particular development forced them to reconsider the status of possible worlds. In the mid-1950s Saul Kripke's interpretation of modal logics in terms of possible worlds gave a boost to various metaphysical speculations about how one should understand these worlds. Three or four main positions seemed to crystalize: (1) possible worlds are all abstract entities except the actual world, (2) possible worlds are all regarded as concrete beings, or (3) possible worlds are taken to have no beings at all, but held to be either fictions or mere ways of speaking. It is clear that an evolutionary naturalist would subscribe to the third position. But which arguments does he hold against the other two?

Here I shall focus on David Lewis' interpretation, in which possible worlds are concrete worlds causally and spatiotemporally separated from each other. The term the "actual world" is just an index, which every human being would use about the world of which he or she takes part. In other worlds different from our actual world, we may have our counterparts that are very similar to us but not quite identical. We do not live in two different worlds, just in the same way as we cannot be in two different places at the same time. The alleged advantage of this interpretation is that we have the truth-makers that seem to endow our modal and counterfactual statements with a truth value. For instance, if I say "It is

[8] Lewis (1986).

possible that I am in Copenhagen tomorrow and not Granada," it means that "There is a least one world in which I am in Copenhagen tomorrow and not Granada" is true.

However, there are numerous problems with Lewis's approach; one of the recognized problems in connection with this interpretation is that "I" in the second sentence is not me, but my counterpart "me," whereas "I" in the first sentences concern one who is identical to me. However, the modal discourse concerning me is about me and not about somebody who is similar to me, no matter how great that similarity might be. It will not improve this interpretation to say that my actual "I" is identical to the other "I" because it does not make much sense to claim that I am identical to somebody who is causally and spatiotemporally separated from me.

Another problem is that since by definition these non-actual possible worlds cannot causally interact with this actual world, it would not stop us from using modal terms (or believing in modal properties) even if these possible worlds do not exist, either as abstract or as concrete beings. Moreover, we do not have an explanation for how it is possible to recognize the truth of an existential claim concerning something about which we can never have any sensory information.

The evolutionary naturalist would argue that the idea of concrete non-actual worlds similar to the actual world is a product of the hypostatization of our reflective thinking. Indeed, the evolution of the capacity of reflective thinking was a huge benefit for *Homo sapiens* in its struggle for survival, but as soon as we use reflective thinking on issues that could not have played any role in the selection of this capacity, we cannot ever claim to know the existence of possible worlds. Instead the evolutionary naturalist will say something like the following: The sentence "There is a possible world in which *x* is *F*" means nothing but "It is possible that *x* is *F*," which again means "*x* is sometimes *F* and sometimes not-*F*." We have noticed many times in our environment that *x* is occasionally *F* and occasionally not-*F*. Hence, saying that it is possible that *x* is *F* is merely an easy way for us to express this discovery. The meaning of possibility statements finds its epistemic support in those cases where we have been empirically able to establish that *x* is sometimes *F* and sometimes not. However, in all other cases where we use the term "possible," this use has been abstracted from these inductively discovered situations.

Yet, this does not take care of the non-actualized possibilities. How do we ever come to believe that some states of affairs are possible, but have in fact never been actualized? We often use words in a metaphorical sense, say, "He is the black sheep of the family." A metaphor is a figure of speech that makes an implicit, implied, or hidden comparison between two things or objects. Similarly, by seeing an analogy between observed situations and some imagined but never observed situations, we have extended the use of "possibility" and "necessity" and use them "metaphorically" to cover these non-actualized possibilities.

The situation is very similar in science. Around the same time that Kripke proposed the possible world semantics of modal logics, Hugh Everrett suggested a new interpretation of quantum mechanics. This interpretation was later named the many-worlds interpretation. Quantum mechanics relies heavily on the use of probabilities, and probability statements may be interpreted as modal expressions. In a brief form the many-worlds interpretation assumes that all possible outcomes of a measurement, expressed in the terms of the wave function, are realized in so many disparate worlds. The claim is an attempt to deal with the so-called measurement problem, according to which a quantum system described by Schrödinger's wave function undergoes a sudden change during a measurement from being in a state of superposition to a state of having a definite value.

As such, probability statements like other modal expressions seem to get their truth value by reference to the possible state of affairs in another possible world different from the actual world. A probability statement like "It is highly probable that I will be in Granada two days from now," uttered two days ago when I was in Copenhagen, would then mean "In most of the possible worlds, in which I (or my counterparts) exist, it is or will be, if these other worlds are tensed, true that I am in Granada two days from now." Apparently, the representational intuition behind the many-worlds interpretation is the same as the one behind the realist interpretation of possible worlds. And just as we saw in the case of modal discourse, the realist commitments may involve either abstract worlds or concrete worlds as the proper truth-makers.

The many-worlds interpretation takes a realist understanding of the probabilities derived from the wave function for granted by claiming that there is a measurement problem, because they assume the mathematical collapse refers to a real event and is not an artifact of the symbolism. They claim to solve this by associating every possible outcome of a particular wave function with a value realized in its own world. Many speculative suggestions attempt to deal with the separation of the actual measured value from all those other values not registered in the actual world. Does the system measured have its counterparts in other concrete worlds or does the actual world physically split into as many worlds as there are possible outcomes? Nobody has been able to give a satisfactory answer that does not suffer from the problem we saw in connection with a realist interpretation of possible worlds or that does not run into absurd divisions of everything connected to the system exposed to a measurement. Even the observer's mind, assuming it is an entity, has to give up its identity because it will also be part of this division. Physicists believing that the splitting of a world involves a spatiotemporal separation of one and the same system into several "branches" have a problem in explaining how the physical mechanism of such a separation works. The mechanism must neither reduce nor increase the system's mass, energy, charge, etc.

Thus, the many-worlders are metaphysical realists. The kind of realism they support is *representationalist* realism, the presumption that certain theoretical models employed to give the model empirical feet on the ground are *representations* of an external reality. This is a premise for the many-worlds view. If one rejects it, there is no reason to postulate any multiverse of branching "worlds." It seems to me that the sheer ontological extravagance of the many-worlds view ought to be sufficient warning that the representationist premise is faulty, but apparently that inference is not appealing to the many-worlds' advocates. Many-worlders are representationalists with a vengeance.

In general, today physicists see quantum mechanics as a theory that applies not only to micro-physical objects, but also to macro-physical objects as well as the entire universe. In certain camps this extrapolation of quantum theory to the macro level has only strengthened the many-worlds interpretation with respect to cosmic inflation and the so-called initial conditions problem. For example, Viatscheslav Mukhanov says, "The discovery of quantum mechanics was in fact the discovery which gave a solid scientific basis to the 'Multiverse versus Universe' debate."[9] Max Tegmark adds, "Accepting quantum mechanics to be universally true means that you should also believe in parallel universes."[10] And Frank Tipler draws the same conclusion: "More precisely, if the other universes and the multiverse do not exist, then quantum mechanics is objectively false. This is not a question of physics. It is a question of mathematics."[11] Despite such supportive remarks in favor of a multiverse, all such claims are entirely without any empirical support. It would not stop us from believing in quantum mechanics if these parallel universes do not exist, but the belief would not be in a realist reading of the collapse of the wave function. Neither is their existence supported by any physical or philosophical arguments. As far as I can see, the assumption that there are parallel universes causally and spatiotemporally separated from ours is based on two highly speculative claims.

The first claim is that the universe is essentially a mathematical structure of which material "particles" and "forces" are in some unexplained way "manifestations." Thus, mathematics is not a descriptive tool invented by us to talk about an external reality. Physical reality is nothing but the existence of a mathematical structure. Such a claim is no more or less empirically based than the common belief in God. The second claim is that the mathematical formulation of quantum mechanics has to be understood quite literally as expressed by the many-worlds interpretation. There is basically no difference between a mathematical formulation of a theory and its physical interpretation. Every mathematical structure emerges as a possible [or perhaps an "actual" (only not this "actual")] physical reality.

Modern Cartesian physicists are of course aware of the fact that mathematical laws require the input of initial conditions. Today these initial conditions can be determined only empirically, but this traditional separation between physical laws and contingent initial conditions was already questioned by Paul Dirac. One way to solve this initial condition problem would be to hold that "everything that exists mathematically is also endowed with physical existence." Though all possible initial

9 Mukhanov (2007), p. 33.

10 Tegmark (2007), p. 23.

11 Tipler (2007), p. 93.

conditions cannot be realized in the same world, each logically possible set of conditions could be realized in infinitely many parallel universes. As Tegmark emphasizes, "All properties of all parallel universes…could in principle be derived by an infinitely intelligent mathematician."[12] A similar message is expressed by Stephen Hawking and Leonard Mlodinow:

> We will describe how M-theory may offer answers to the question of creation. According to M-theory, ours is not the only universe. Instead, M-theory predicts that a great many universes were *created out of nothing*. Their creation does not require the intervention of some supernatural being or god. Rather, these multiple universes *arise naturally from physical laws*. They are a prediction of science.[13]

It should be noted, as a curiosity, that these multiple universes are alleged to arise physically out of nothing as a manifestation of purely mathematical laws of nature. What does this really mean? In my opinion, the very idea that the actual universe, together with many other universes, was created out of nothing shows that basic cosmology has reached the cognitive limits of doing proper science.

10.4 The Copenhagen Interpretation: A Non-Representational View

The basic premise of the many-worlds interpretation is that the state vector represents the physical state of a system. The many-worlders are not the only realists to assign an ontological status to the state vector, but these other realist proposals such as configuration space realism, wave-function realism, and dispositionalism suffer from the same calamities as many-worlds realism in that they postulate the existence of a host of causally inert or abstract structures.[14] One consequence of wave-function realism is that the observation of a quantum system involves a process in

[12] Tegmark (2008), p. 125.

[13] Hawking & Mlodinow (2010), p. 8. Italics mine.

[14] The reader may consult Ney and Albert (2013) in which various authors outline various ontological interpretations of the wave function. An excellent and painstaking discussion of realism versus

which the evolution of the state vector in configuration space undergoes an acausal transition from a real superposition of eigenstates of measurable observables to the specific eigenstate corresponding to the measured value in perceptual space.[15] This is the infamous measurement problem. However, it is also clear that it arises as a problem only because of the representational approach to laws of nature, and specifically to Schrödinger's wave function. Within such an approach one can continue along three different lines by arguing: (1) Measurements are physically very different from microscopic processes in virtue of decoherence; (2) the quantum mechanical formalism is incomplete and must be supplemented with hidden variables, or (3) the quantum mechanical formalism is complete but reality does not undergo an acausal transition. Instead every eigenstate represented in a superposition by the state vector is 'real;' no unobserved value is any less 'real' than the value of the observable measured in the actual world.

Fortunately, the non-representationalists are not without allies. Much closer to a naturalist point of view is Niels Bohr's pragmatic interpretation of quantum mechanics. Although there are no records of Bohr specifically writing or speaking about Darwin and human evolution, he strongly endorsed the position that the interpretation of quantum mechanics should be confined to what we possibly can experience. He often mentioned the unavoidable use of classical concepts, which he regarded as necessary for the description of ordinary human experience, because human beings are adapted to understand their physical environment in terms of these concepts. As the basic cognitive means by which we grasp our perceptual experience, classical concepts became embedded in our natural language long before the discoveries of classical physics. For instance, he wrote:

> From a logical standpoint, we can by objective description only understand a communication of experience to others by means of a language which does not admit ambiguity as regard the perception of such communica-

---

instrumentalism concerning the wave function is to be found in Dorato & Laudisa (2014) and Dorato (2015).

[15] The alleged "collapse" in some interpretations is imagined to be a mental-physical interaction. The mind's intervention into the physical order is conceived as the "cause" of the alleged collapse.

> tions. In classical physics, this goal was secured by the circumstance that, apart from unessential conventions of terminology, the description is based on pictures and ideas embodied in common language, *adapted to our orientation in daily-life events*.[16]

Or again: "All account of physical experience is, of course, ultimately based on common language, *adapted to orientation in our surroundings and tracing of relationships between causes and effects*."[17] So in a manner similar to an evolutionary naturalist, Bohr seemed to have believed that it is possible to understand quantum mechanics only if we take it to be a theory that can tell us something about the world as it appears in everyday human experience. If we attempt to say something about how the world-is-in-itself we exceed the reach of our empirical concepts.

Typically pragmatism rejects the traditional Cartesian assumption that the function of thought is to depict, mirror, or represent reality. A pragmatist believes that mental activity is a result of an interaction between the organism and its environment, not only in the individual organism but also in the species over evolutionary time. Thus, thoughts and later language (including mathematics) are first and foremost created as a capacity for handling external information in order for the organism to solve problems. Solving problems is paramount: Predicting and explaining are steps toward solving problems. So believing that mathematics represents the world as it is in itself indicates that one has fallen victim to the Cartesian representationalist view of mathematical knowledge.

Instead, the pragmatist would argue that scientific theories function as tools for predictions and explanations, and any theory should be evaluated according to how successful it is in fulfilling these aims. Scientific theories work as tools because they have been created as a conceptual systematization of our scientific observations (just as our common sense conception of the world allows us to systematize our everyday observa-

[16] Bohr (1998), pp.156–157. Italics mine. See also Bohr (1998), p. 176. Here he opposed the view that the present quantum mechanics could be replaced by a deterministic theory, while pointing out that the "vivid discussion of this basic issue has greatly stimulated the analysis of our position as observers of nature and especially stressed the caution necessary in the application to a new domain of knowledge of concepts *adapted to our orientation under ordinary conditions*." Italics mine.

[17] Bohr (1963), p. 1. Italics mine.

tions). The success of a theory should be measured with respect to its ability to describe these observations accurately and not with respect to claims that cannot be experimentally tested. The pragmatist would say that any interpretation of a scientific theory that assumes it represents the world ignores the fact that knowledge begins and ends in experience.

A very similar pragmatist attitude is to be found in Bohr: He says, for instance, that the purpose of scientific theories "is not to disclose the real essence of phenomena but only to track down, so far as it is possible, relations between the manifold aspects of experience."[18] We do this by seeing how well our theories are able to predict experimental results accurately. In other places he distances himself from the representationalist view of theories: "[T]he ingenious formalism of quantum mechanics, which abandons pictorial representation and aims directly at a statistical account of quantum processes …"[19] Similarly, he says "The formalism thus defies pictorial representation and aims directly at prediction of observations appearing under well-defined conditions."[20] Bohr's interpretation of quantum mechanics shares with the pragmatists a clear-cut rejection of theories as mathematical representations of the inner structure the world.

Indeed, this pragmatic approach to scientific theories had consequences for Bohr's view of the quantum state. First of all, the wave function should not be interpreted in any pictorial sense, or rather *cannot* be so interpret because it is a function in a multidimensional Hilbert space. The state vector has only a symbolic and calculative function. It does not represent anything; it does not represent some kind of novel "quantum reality," but "[t]he entire formalism is to be considered as a tool for deriving predictions of definite and statistical character … "[21] One of Bohr's reasons for considering the wave function to be symbolic is that Schrödinger's wave function is defined in terms of imaginary numbers. Real numbers can be associated with the recording of measurement values in ordinary space and time, whereas imaginary numbers function only in abstract vector spaces and have no counterparts in measurement.

---

[18] Bohr (1958), p. 71.

[19] Bohr (1998), p. 152.

[20] Ibid., p. 172.

[21] Ibid., p. 144.

However, Bohr had two more equally important reasons to regard the wave function as symbolic: First, its mathematical interpretation is formulated in a configuration space that usually contains more than three dimensions, and second the deterministic evolution of the wave function is not limited by the speed of light.

All this indicates that Bohr did not think of the measurement as a "physical" collapse of the wave function. According to him, the reduction of the wave function that takes place when an experimental outcome yields a definite result does not represent a physical process, and therefore according to his lights there is no measurement problem in quantum mechanics.

The measurement problem is a problem only for those physicists who assume that the wave function must be understood as corresponding to a real quantum state and that the quantum state consists of a complex combination of all possible measurable values. These realists see the Schrödinger wave equation as a literal description of the continuous evolution of the quantum system as a superposition of different states. This superposition of the quantum state represented by the continuous development of the wave function abruptly collapses when the system interacts with a second system—be it a particle or a measuring device—and the system manifests a particular value for a measured parameter. For those who make this assumption, the wave function does not solely represent the probability of possible experimental outcomes, but an actual, existing complex quantum state of the system. Once this assumption is made, then, of course, some physical mechanism "selecting" a particular value during the process of measurement must be at work. (Unless, one takes the many-worlds path, in which case nothing is selected, everything possible happens.) The GRW interpretation is an attempt to describe such a mechanism.

For his part, Bohr understood the wave function as a probability amplitude, and he followed Max Born in his proposal that the modulus squared of this quantity should be interpreted as a probability density. Bohr believed that the Born probability is only an expression of how likely it is to *measure*, say, a certain momentum or a certain position of a quantum system. It cannot, he thought, be interpreted ontologically as a probability density for a quantum system in *being* at a certain place or *having* a certain momentum independently of its entering into an interaction with a measuring system. If an ontological interpretation of the Born probability rule had been possible at all, it would have made sense to say that the same rule could be used for ascribing a certain probability to a quantum object for going through one or the other slit in the double-slit experiment. But in those circumstances such an ascription of probabilities to "which ways" is not an element of standard quantum mechanics. Instead, I shall argue below that Bohr took the probability of *observing* a quantum system at a certain place to be equivalent to the probability of *assigning* the system that very position.

Nothing in Bohr's writings indicates that he excluded the possibility of using the quantum mechanical formalism to describe macroscopic systems or parts thereof. When he does talk about the distinction between classical and quantum mechanical descriptions, he is careful to point out that it is not an absolute distinction but one that depends on the experimental situation. It would also have been inconsistent for him to advocate an absolute distinction since he did not believe in the representational nature of our scientific theories. So he would admit that even macroscopic objects can be treated as quantum objects, though ones of enormous complexity. Nonetheless, he insisted on the use of classical concepts in the description of the outcome of physical experiments and in the account of the results that were produced by these experiments. By "classical" concepts he meant first and foremost 'space,' 'time,' 'position,' 'duration,' 'momentum,' and 'energy.' As he said: "However far the phenomena transcend the classical physical explanation, the account of all evidence must be expressed in classical terms."[22] Why did he hold this view?

Bohr had at least three arguments for the indispensable use of classical concepts. The first results from his adoption of the correspondence principle, which stipulates that quantum mechanical predictions for high numbers approach classical predictions. In the beginning this methodological principle seems to have had a structural formulation, but it is clear from Bohr's writings that he eventually realized that in order to work, the principle also had to have a semantic dimension built into it. One cannot compare predictions of two theories unless these two theories describe their observational evidence in the same vocabulary. Hence, classical concepts are essential for our understanding why perhaps

[22] Bohr (1958), p. 39.

one physical theory fails and another is successful. The communication of scientific knowledge of the physical world is based on these concepts, and those same concepts are embedded in any understanding of the practice of the physicists' quest for knowledge. So whatever theory one may propose, it has to use the classical concepts to describe empirical results.

The second argument begins with the possibility of measurement. In Bohr's own formulation:

> … in each case some ultimate measuring instruments, like the scales and clocks which determine the frame of space-time coordination—on which, in the last resort, even the definitions of momentum and energy quantities rest—must always be described entirely on classical lines, and consequently kept outside the system subject to quantum mechanical treatment.[23]

Every physical experiment has to relate to a spatially and temporally fixed coordinate system in order for the physicists to be able to identify and determine measurable properties. It is only in relation to a spatially and temporally fixed frame of reference that physicists can define a body as moving or at rest. Hence, it is only with respect to an empirically established frame of reference that position and momentum are well defined. Acting as frames of reference, the measurement experiments must have a classically well-defined position and a classically well-defined momentum in relation to classically defined space and time in order for the outcomes of the experiments to be communicable as unambiguous and determinate results. In such circumstances the properly described measuring instrument can serve to define the frame of reference for "those parts which are to be regarded as objects under investigation and in the account of which quantum effect cannot be disregarded."[24] Thus, classical concepts are indispensable. It is only through their use that it is possible to establish a frame of reference in terms of which instrument recordings make sense.

The third argument is in some sense the most fundamental. Bohr's thinking seems to be that from very early in our evolution experiential concepts of human beings have been adapted to their environment in

[23] Bohr (1998), p. 104.

[24] Bohr (1949), p. 228.

order for us to understand our perceptions. This is essentially the pragmatists' view. Originally these concepts specified a determinate position of an object in relation to other objects and a determinate spatial distance in between them, as well as the simultaneous occurrence of enduring events or a temporal separation between them, and particular events causing other events. Much later these basic concepts became fundamental categories in our natural language, and even later physicists were able to develop a mechanical physics based on a definitional precision of these concepts. Hence, because of the fundamental status of these concepts for describing our experience unambiguously, we cannot unequivocally communicate our experience if it is not described in those categories. If we take this suggestion a bit further, we may say that our common language reflects genetically based conceptual schemata of perceiving in our brains.

So Bohr concluded that the classical concepts are obligatory for understanding quantum mechanics. But what happened from the transition of classical physics to quantum physics was that the use of classical concepts became restricted to the outcomes of particular experimental situations. It was this restriction that gave rise to his famous view of complementarity.

Bohr embraced objectivity but not traditional scientific realism. He was an entity realist, as this label is used by Ian Hacking and Nancy Cartwright, but not a theory realist. Even if he denies semantic realism, as I have emphasized, he does admit the reality of the entities, so there is some "external world" commitment here. More than once in his writings Bohr connected a pragmatic notion of objectivity with unambiguous communication. "Every scientist is constantly confronted with the problem of objective description of experience, by which we mean unambiguous communication."[25] Again this is a feature of pragmatism. Our interaction with the world tells us what is objectively real. Action and its reflective prolongation, the scientific experiment, are the cognitive means human beings—including physicists—have for determining what is real. Neither passive sense perception nor mathematical calculation can tell human beings what is real. The only situation in which we can speak unambiguously about atomic objects is in connection with experiments.

[25] Bohr (1958), p. 67.

> The argument is simply that by the word 'experiment' we refer to a situation where we can tell others what we have done and what we have learned and that, therefore, the account of the experimental arrangement and of the results of observations must be expressed in unambiguous language with suitable application of the terminology of classical physics.[26]

It is only by interaction with things that we can claim that we have empirical knowledge; these things must be real in order for our perceptual organs to interact with them and do not suffer from a perceptual or theoretical illusion. Hence, as a good scientist with a pragmatic outlook, Bohr would say that we understand the world in particular through our actions and experiments and that our knowledge of the world does not go beyond the information we gain by these actions and experiments, information that can be formulated in common sense language supplemented with a technical terminology. Scientific theories help us in predicting what will happen whenever we carry out these actions and experiments.

We use theories to construct models, and employing these theoretical models the theoretician is able to make empirical predictions, but that does not, of course, imply that the model is a *representation* of an external reality in itself. The empirical predictions always refer to that reality in an interaction with the physical system used to obtain the empirical data. This is one of Bohr's main points, ignored as negligible or correctable by classical mechanism. However, the abandoning of the conceit that neither the theory nor its models are "representations" of reality in itself does not amount to forsaking realism. The entity realism, which I think characterized Bohr's view, is based on the same sort of inductive inferences as those by which we establish the existence of ordinary object. Of course these inferences are based on an indemonstrable assumption of a causal connection between the world of human experience and an external world. This is the view I not only associate with pragmatism. One might also seek to mount a naturalist defense that natural selection has hard-wired this assumption into our immediate consciousness of a world external to ourselves and indeed that natural selection has arrived at this strategy for survival because it represents how things really and truly are for us.

[26] Ibid., p. 39.

11

Conclusion

Taking our biological heritage seriously leads us to conclude that our cognitive capacities have evolved through variation, adaptation, and inheritance. If we look at how human beings and chimpanzees appear, we are all physically very similar. But we are also very similar when it comes to most of our basic cognitive abilities. Both species are able to process sensory information from our environment so that this information gives rise to knowledge of the environment, i.e., both species process information by which they acquire usually true beliefs that are reliably produced (perceptual illusions are possible, but not the usual order of the day). Moreover, based on earlier sensory information, both species can imagine how the near future will turn out to be, and they can also imagine how their actions may intervene in the course of events. And they can learn which actions they benefit from under given circumstances. Hence, based on inductive inference, both species act according to a strategic plan.

There are undeniable differences between the two species, too. Humans make use of highly developed languages and mathematical theories, whereas chimpanzees do not have any of these capacities. Still chimps use non-linguistic concepts, and they can count and calculate small numbers. So much of the required capacity for the evolution of language and

J. Faye, *Experience and Beyond*,
DOI 10.1007/978-3-319-31077-0_11

the development of mathematics was already in place in our common ancestors. However, we can reasonably imagine that a more advanced system of reflective thoughts and behavioral communication evolved in hominids before any spoken language evolved in *Homo sapiens*. But in virtue of the rise of an oral language, *Homo sapiens* got a new and powerful ability to learn from and cooperate with other members of their own species. The way was open for the development of science and modern civil society.

Darwinians tell us that our specific adaptations are in principle accidental. That our main sensory inputs are visual stimuli in a rather narrow band of electromagnetic frequencies is one of several ways evolution could have taken. Our sensory system could have evolved differently as we know from bats' use of acoustic echolocation. But the evolution of any new cognitive mechanism is always determined by the environment, as this can be physically presented to the organism in virtue of some sensory channels. The consequence is that the cognitive mechanisms by which we gain knowledge of the external world are fitted by natural selection to processing the information we received through our visual senses. The scientific details are complex, and cognitive neuroscience and neuropsychology are just at a relatively early stage of development. But at least this much can be said: Biological evolution has adapted our brains to handle sensory-based information about the environment mainly in relation to our possible behavioral performances. One part of this adaptation is the visual ability to separate and individuate objects distinct from the visual background upon which they are situated. Another part is our ability to recognize types of objects. We must first have a notion of similarity—and one that admits of degree: a horse and a zebra are similar, but not as similar as two horses are to each other. So we can identify other members of our own species and different species. In some sense we see the similarity as directly as, say, the color. All of this has to come at a quite early stage of evolution. If an organism is going to learn from the past and be able to tackle similar situations in the future, it must be able to recognize the most important objects in its environment as of the same type as those experienced in the past in order to use what it already has learned. Thus, the ability to abstract concepts and to receive knowledge from visual information seems to be part of our biological heritage.

It gives those organisms that have such skills the capacity to think and reflect about what they are sensing and how to react to it in a manner that enhances the reproductive success of the species.

Apparently another capacity evolved with hominids that may be found in only a few other animals. This is the capacity to reflect upon one's own images and thoughts. This capacity has given those who possess it a better chance of surviving to reproduce than any other beast. By possessing such a capacity they could reconsider their wishes and intentionally correct their images and thoughts if they led to unwanted consequences. In order for reflection to work we need to have an understanding that our wishes, perceptions, imaginations, and thinking *stand for* something that is different from our wishes, perceptions, imaginations, and thinking themselves. So the evolution of reflective thinking gave hominids huge advantages but also some disadvantages.

The rise of reflective thinking and the ability to represent their mental content created in human beings a strong psychological tendency to hypostasize concepts to stand for a reality divided from the concrete items from which these concepts originally were abstracted. The reality corresponding to these abstracted concepts is either seen as a realm of Platonic objects or as a realm of possible worlds that are causally and spatiotemporally separated from our world. But getting to know such alleged realities, if they are real, is not something to which we are cognitively adapted. In my opinion, philosophers and physicists following the innate tendency of reification ignore what kind of knowledge reflective thinking cannot provide, namely knowledge about subjects to which our cognitive powers have not been biologically adjusted to grasp. We cannot know what selection has not made possible for us to know.

It is not unreasonable to argue that among our cognitive dispositions it is impossible to have a capacity of gathering information about some alleged reality that is defined to be causally isolated from our world. Such other allegedly real worlds could not have had any influence on the evolution of our mental capacities. Consequently, evolution sets limits to what we can possibly know and therefore what science and philosophy can legitimately claim to know.

Much debate about the nature of reality revolves around two extreme and implausible positions. One is strong metaphysical skepticism claiming that whatever exists is *merely* a result of our imagination and psychological habits. The other is strong metaphysical realism holding that the nature of reality may transcend far beyond our cognitive limitation of knowledge such that its nature may be quite different from what we can get to know by experiments and observations. The evolutionary naturalist is neither a metaphysical skeptic nor a metaphysical realist in any of these extreme senses. She is rather an agnostic metaphysical antirealist. We are born with a realist instinct that is due to our biological adaptation to an external environment, but the adjustment of this very instinct is also confined to the physical reality that caused its selection.

Once Frank Ramsey remarked, "When a philosophical dispute presents itself as an irresolvable oscillation between two alternatives, the likelihood is that both alternatives are false and share a common false presupposition."[1] Since the prolonged debate between scientific realists and empiricists about laws of nature and theoretical entities appears to be such an "irresolvable oscillation" between metaphysical realism and metaphysical antirealism, we should appreciate Ramsay's insight and look for the common ground that led both views astray. We have argued that both scientific realism and empiricism build and feed on the dualism between a realm of visibles and a realm of invisibles, between the notion of how things appear as visual presentations and how things really are. Thus, on the one hand, the realist claims that the world-in-itself is as it is quite independently of our cognitive powers and thus that we can never be certain that we really know it. As Michael Devitt says, "Realism's independence dimension makes it possible that we could be completely wrong about the world. The principle of charity denies that we could: For the most part, if we believe that *p*, then *p*."[2] The evolutionary naturalist argues that although individual belief could be false, blanket skepticism is not possible. The constructive empiricist, on the other hand, holds that we cannot be certain about what exists behind the immediate sense experience, adding that because we can at least apply the principle of charity to visible things, we should limit what is true to what can be perceived with the naked eye and confine justified true statements to sentences

[1] Ramsey ([1930]1990).
[2] Devitt ([1984]1991), p. 199.

concerning empirical phenomena only. In contrast, the evolutionary naturalist would argue that knowledge about the world is not limited to experience delivered by our sense organs but can be extended to include information from our technological devices and instrumental recordings. Thus, realists are in accord with evolutionary theory in so far as they claim that there exists a physical world inaccessible to the naked eye. However, the evolutionary naturalist also argues that our concepts of things with which we cannot causally interact refer only to abstractions created by the mind. So empiricists are correct insofar as they hold that numbers, concepts, universals, or any form of possible worlds do not exist independently of human thought.

Bibliography

Alexander, H.G. 1956. *The Leibniz-Clarke Correspondence*. Manchester: Manchester University Press.

Aristoteles. 1955. *Physics*, The W.D. Ross-edition edn. Oxford: Clarendon Press.

Armstrong, David M. 2004. *Truth and Truth-makers*. Cambridge: Cambridge University Press.

Amundson, Ron. 1989. The Trials and Tribulations of Selectionist Explanations. In *Issues in Evolutionary Epistemology*, eds. K. Hahlweg and C.A. Hooker. New York: SUNY Press.

Austin, John L. 1950. Truth. Reprinted in G. Pitcher (ed.) Truth, Englewood Cliffs: Prentice-Hall, 1964, 18–31.

Bacon, John. 1995. *Universals and Property Instances: The Alphabet of Being*, (Series: *Aristotelean Society Series*. Vol. 15). Oxford: Blackwell.

Baker, Alan. 2005. Are there Genuine Mathematical Explanations of Physical Phenomena? *Mind* 114: 223–236.

Belot, Gordon, and John Earman. 1999. From Metaphysics to Physics. In *From Physics to Philosophy*, eds. J. Butterfield and C. Pagonis, 166–186. Cambridge: Cambridge University Press.

———. 2001. Pre-Socratic Quantum Gravity. In *Physics Meet Philosophy on the Planck Scale*, eds. C. Callender and N. Huggett, 213–255. Cambridge: Cambridge University Press.

J. Faye, *Experience and Beyond*,
DOI 10.1007/978-3-319-31077-0

Blackburn, Simon. 1993. Hume and Thick Connexions. In *In his Essays in Quasi-Realism*. Oxford: Oxford University Press.

Bohr, Niels. [1920]1976. On the Spectra of the Elements. In *Collected Works*, vol 3, ed. Leon Rosenfeld. Amsterdam: North Holland.

———. [1925]1984. Atomic Theory and Mechanics. In *Collected Works*, vol 6, ed. Leon Rosenfeld. Amsterdam: North Holland.

———. 1949. Discussion with Einstein on Epistemological Problems in Atomic Physics. In *Albert Einstein: Philosophy and Scientist*, (Series: *The Library of Living Philosophers*. Vol. 7)., ed. P.A. Schilpp, 199–243 . La Salle: Open Court.Reprinted in *Atomic Physics and Human Knowledge*, 32–66.

———. 1958. *Atomic Physics and Human Knowledge*. New York: John Wiley & Sons.

———. 1963. *Essays 1958–1962 on Atomic Physics and Human Knowledge*. New York: John Wiley & Sons.

———. 1998. *Causality and Complementary. Niels Bohr's Philosophical Writings.* 4. Edited by J. Faye & H. Folse. Woodbridge, CT: Ox Bow Press.

Bourdieu, Pierre. [1997]2000. *Pascalian Meditations*. Stanford: Stanford University Press.

Boyd, R.N. 1973. Realism, Underdetermination, and a Causal Theory of Evidence. *Nous* 7: 1–12.

———. 1985a. *Lex Orandi est Lex Credendi*. In *Images of Science*, eds. P.M. Churchland and C.A. Hooker. Chicago: The University of Chicago Press.

———. 1985b. The Current Status of Scientific Realism. In *Scientific Realism*, ed. Jarrett Leplin. Berkeley: University of California Press.

———. 1990. Realism, Approximate Truth, and Philosophical Method. In *Scientific Theories*, Series: *Minnesota Studies in the Philosophy of Science*, vol. 14, ed. S.W. Savage. Minneapolis: University of Minnesota Press.

Boysen, S.T. 1993. Counting in chimpanzees: Nonhuman principles and emergent properties of number. In *The Development of Numerical Competence*, eds. S.T. Boysen and E.J. Capaldi, 39–59. Hillsdale, NJ: Lawrence Erlbaum Associates.

Boysen, S.T., and K.I. Hallberg. 2000. Primate Numerical Competence: Contributions Toward Understanding Nonhuman Cognition. *Cognitive Science* 24: 423–443.

Bradie, Michael. 1989. Evolutionary Epistemology as Naturalized Epistemology. In *Issues in Evolutionary Epistemology*, eds. K. Hahlweg and C.A. Hooker. New York: SUNY Press.

Brown, Harold I. 1987. *Observation and Objectivity*. New York: Oxford University Press.

Brown, Michael F. 2006. Abstracting Spatial Relations among Goal Locations. In *Animal Spatial Cognition: Comparative, Neural and Computational Approaches*, eds. M.F. Brown and R.C. Cook.Available: www.pigeon.psy.tufts.edu/asc/

Burgess, J. and G. Rosen. 1997. *A Subject with No Object*. Oxford: Clarendon Press. Available: www.pigeon.psy.tufts.edu/asc/

Butterfield, Jeromy. 1989. Albert Einstein Meets David Lewis. In A. Fine & J. Leplin (eds.) *PSA Proceedings of the Biennial Meeting of the Philosophy of Science Association, 1988*, Vol. 2, 65–81.

Cameron, Ross P. 2008. Truthmakers, Realism and Ontology. In R. LePoindevin (ed.) *Being: Contemporary Developments in Metaphysics*. Cambridge: Cambridge University Press, 107–128.

———. 2010. How to Have a Radically Minimal Ontology. *Philosophical Studies* 151: 249–264.

Campbell, Donald T. 1974. Evolutionary Epistemology. In P.A. Schlipp (ed.) *The Philosophy of Karl Popper*. Vol. 1. (Series: Library of Living Philosophers). La Salle: Open Court, 413–459.

Campbell, Donald T., and Bonnie T. Paller. 1989. Extending Evolutionary Epistemology to Justifying Scientific Beliefs (A Sociological Rapprochement with a Fallibilist Perceptual Foundationalism?). In *Issues in Evolutionary Epistemology*, eds. K. Hahlweg and C.A. Hooker. New York: SUNY Press.

Campbell, Keith. 1981. Abstract Particulars. In *Midwest Studies in Philosophy*. Reprinted in D.H. Mellor & A. Olivier (eds.) *Properties*. Oxford: Oxford University Press, 1997, 125–139.

Cartwright, Nancy. 1989. *Nature's Capacities and their Measurement*. Oxford: Oxford University Press.

Carnap, Rudolf. [1950]1958. Empiricism, Semantics, and Ontology. In *Revue internationale philosophie*, 4. Reprinted in the second edition of *Meaning and Necessity*. Chicago: The University of Chicago Press.

Chisholm, Rodrick. 1982. *The Foundation of Knowing*. Minneapolis: University of Minnesota Press.

Chomsky, Noam. 2004. "Language and Mind: Current thoughts on ancient problems", Part I & Part II. In *Variation and Universals in Biolinguistics*, ed. Lyle Jenkins, 379–405. Amsterdam: Elsevier.

———. 2005. "Three factors in language design. *Linguistic Inquiry* 36 (1): 1–22.

Church, Russell M. 2002. Temporal learning. In *Stevens' Handbook of Experimental Psychology*, Learning, Motivation, and Emotion, vol 3, 3rd edn, eds. H. Pashler and R. Gallistel, 365–393. New York: Wiley.

———. 2003. Time Perception. In *Encyclopaedia of Cognitive Science*, ed. L. Nadel. New York: Macmillan.

Craig, William. 1956. Replacement of Auxiliary Expressions. *Philosophical Review* 65: 38–55.

Cruse, Pierre, and David Papineau. 2002. Scientific Realism Without Reference. In *The Problem of Realism*, ed. M. Marsonet. Aldershot: Ashgate Publishing Company.

Davidson, Donald. 1969. "The Individuation of Events". In N. Rescher (ed.) *Essays in Honor of Carl G. Hempel*. Dordrecht: Reidel, 368–274.

Devitt, Michael. [1984]1991. *Realism & Truth*. Cambridge, MA: Blackwell.

DiSalle, Robert. 1994. On Dynamics, Indiscernibility, and Spacetime Ontology. *British Journal for the Philosophy of Science* 45: 265–287.

———. 2002. Conventionalism and Modern Physics. A Re-Assessment. *Nous* 36(2): 169–200.

Dorato, Mauro. 2000. Substantivalism, Relationism, and Structural Spacetime Realism. *Foundations of Physics* 30(10): 1605–1628.

———. 2015. Laws of Nature and the Reality of the Wave Function. *Synthese* 192(10): 3179–3201.

Dorato, Mauro, and Frederico Laudisa. 2014. Realism and Instrumentalism About the Wave Function. How Should We Choose? In *Protective Measurements and Quantum Reality: Toward a New Understanding of Quantum Mechanics*, ed. G. Shan. Cambridge: Cambridge University Press.

Dorato, Mauro, and Massimo Pauri. 2006. Holism and structuralism in classical and quantum general relativity. In *The Structural Foundations of Quantum Gravity*, eds. S. French, D. Rickles, and J. Saatsi, 121–151. Oxford: Oxford University Press.

Dretske, Fred I. 1963. *Seeing and Knowing*. Chicago: University of Chicago Press.

———. 1981. *Knowledge and the Flow of Information*. Cambridge, MA: MIT Press.

———. 2000. *Perception, Knowledge, and Belief*. Cambridge: Cambridge University Press.

Dummett, Michael. 1973. *Frege Philosophy of Language*. Duckworth: London.

———. 1976. What is a Theory of Meaning II? In *Truth and meaning: Essays in Semantics*, eds. G. Evans and J. McDowell. Oxford: Clarendon Press.

———. 1978. *Truth and Other Enigmas*. London: Duckworth.

Earman, John. 1989. *World Enough and Space-Time*. Boston: MIT Press.

Earman, John, and John D. Norton. 1987. What Price Substantivalism? The Hole Story. *British Journal for the Philosophy of Science* 38: 515–525.

Einstein, Albert. 1949. Autobiographical Notes. In *Albert Einstein, Philosopher-Scientist*, (Series: *The Library of Living Philosophers*. Vol. 7), ed. P.A. Schil pp, 1–95. La Salle: Open Court.

———. 1955. *The Meaning of Relativity*. Princeton: Princeton University Press.

Faye, Jan. 1989. *The Reality of the Future. An Essay on Time, Causation, and Backward Causation*. Odense: Odense University Press.

———. 1991. *Niels Bohr: His Heritage and Legacy. An Anti-realist View of Quantum Mechanics*, (Series: *Science and Philosophy.* Vol. 6). Dordrecht: Kluwer Academic Publisher.

———. 2000. Observing the Unobservable. In *The Reality of the Unobservable. Observability, Unobservability and their Impact on the Issue of Scientific Realism*, (Series: *Boston Studies in the Philosophy of Science.* Vol. 215), eds. E. Agazzi and M. Pauri, 165–177. Dordrecht: Kluwer Academic Publisher.

———. 1994. Causal Beliefs and their Justification. In *Logic and Causal Reasoning*, eds. J. Faye, U. Scheffler, and M. Urchs. Berlin: Akademie Verlag.

———. 1997. Is the Mark Method Time-Dependent? and Causation, Reversibility, and the Direction of Time. In *Perspectives on Time*, (Series: *Boston Studies in the Philosophy of Science*. Vol. 157), eds. Jan Faye, Uwe Scheffler, and Max Urchs. Dordrecht: Kluwer Academic Publisher.

———. 2002. *Rethinking Science: A Philosophical Introduction to the Unity of Science*. Aldershot: Ashgate Publishing Company.

———. 2005. How Nature Makes Sense. In *Nature's Principles*, (Series: *Logic, Epistemology, and the Unity of Science*. Vol. 4), eds. J. Faye, P. Needham, U. Scheffler, and M. Urchs, 77–102. Dordrecht: Springer Press.

———. 2006a. Models, Theories, and Language. In *Essays in honor of Evandro Agazzi. Presidenza del Consiglio dei Ministri*, eds. D. Conci, M. Marsonet, and F. Minazzi. Rome: Poligrafico e Zecca dello Stato.

———. 2006b. Science and Reality. In *The Way through Science and Philosophy: Essays in honour of Stig Andur Pedersen*, eds. H.B. Andersen, F. Voetmann Christiansen, V. Hendricks, and K. Frovin Jørgensen, 137–170. London: College Publications.

———. 2006c. Is Time an Abstract Entity? In *Time and History*, (Series) *Proceedings of the 28 International Ludwig Wittgenstein Symposium*, vol *2005*, eds. F. Stadler and M. Stöltzner, 85–100. Ontos Verlag: Frankfurt.

———. 2012. *After Postmodernism. A Naturalistic Reconstruction of the Humanities*. Palgrave Macmillan: Basingstoke.

———. 2014. *The Nature of Scientific Thinking. On Interpretation, Explanation, and Understanding*. Basingstoke: Palgrave Macmillan.

Faye, Jan, Uwe Scheffler, and Max Urchs. 2000. Philosophical Entities. In *Things, Facts and Events*, (Series: Poznan Studies in the Philosophy of the Sciences and the Humanities. Vol. 76)., eds. J. Faye, U. Scheffler, and M. Urchs. Amsterdam: Rodopi.

Field, Hartry H. 1980. *Science Without Numbers. A Defence of Nominalism*. Princeton: Princeton University Press.

Franklin, Anna, Ally Clifford, Emma Williamson, Ian Davies. 2005. Color Term Knowledge Does Not Affect Categorical Perception of Colors in Toddlers. *Journal of Experimental Child Psychology* 90(2): 119–141.

Franklin, Anna, G. V. Drivonikou, L. Bevis, I. R. L. Davies, P. Kay, and T. Regier. 2008. "Categorical Perception of Color is Lateralized to the Right Hemisphere in Infants, But to the Left Hemisphere in Adults." *Proceedings of the National Academy of Science* U.S.A. *105*: 3221–3225.

Friedman, Michael. 1983. *Foundation of Space-Time Theories*. Princeton: Princeton University Press.

Friggs, Roman. 2002. Models and Representation. Why Structures are not Enough In *Measurement in Physics and Economics Project Discussion Paper Series*, DP MEAS 25/02, London School of Economics.

Gärdenfors, Peter. 2003. *How Homo Became Sapiens. On the Evolution of Thinking*. Oxford: Oxford University Press.

Gassendi, Pierre. 1972. *Selected Works of Pierre Gassendi*. Edited and translated by Craig B. Brush. New York: Johnson Reprint Corporation.

Giere, Ronald N. 1988. *Explaining Science. A Cognitive Approach*. Chicago: The University of Chicago Press.

———. 1999. *Science Without Laws*. Chicago: The University of Chicago Press.

Hacking, Ian. 1983. *Representing and Intervening*. Cambridge: Cambridge University Press.

Hanson, Norwood R. 1958. *Patterns of Discovery*. Cambridge: Cambridge University Press.

Hare, B.J., J. Call, and M. Tomasello. 2001. Do Chimpanzees Know What Conspecifics Know? *Animal Behaviour* 61: 139–151.

Hauser, M.D., N. Chomsky, and W.T. Fitch. 2002. The Faculty of Language: What Is It, Who Has It, and How Did It Evolve? *Science* 298: 1569–1579.

Hawkings, Stephen, and Leonard Mlodinow. 2010. *The Grand Design*. New York: Bantam Books.

Hempel, Carl G. [1958]1965. The Theoreticians's Dilemma: A Study in the Logic of theory construction. In *Aspects of Scientific Explanation*, ed. Carl G. Hempel. New York: The Free Press.

Hoefer, Carl. 2000. Energy Conservation in GTR. *Studies in History and Philosophy of Modern Physics* 31(2): 187–199.

Hofweber, Thomas. 2007. Innocent Sentences and their Metaphysical Loaded Counterparts. *Philosophers' Index* 7: 7–33.

———. 2015. Carnap's Big Idea. In *Ontology after Carnap*, eds. S. Blatti and S. Lapointe. Oxford: Oxford University Press.

Horwich, Paul. 1990. *Truth*. Oxford: Basil Blackwell.

Hume, David. 1740. *Treatise of Human Nature*. Ed. Selby Bigge. Oxford: Clarendon Press.

———. 1748. *Enquiry Concerning Human Understanding*. Ed. Selby Bigge. Oxford: Clarendon Press.

Kant, Immanuel. [1787]2007. *The Critique of the Pure Reason*. Translated by Norman Kemp Smith. London: Palgrave-Macmillan.

Kirkham, R.L. 1992. *Theories of Truth. A Critical Introduction*. Cambridge, MA: MIT Press.

Kornblith, Hilary. 1985. Introduction. In *Naturalized Epistemology*, ed. Hilary Kornblith. Cambridge, MA: MIT Press.

Kragh, Helge. 1979. Niels Bohr's Second Atomic Theory. In *Historial Studies in the Physical Sciences*, eds. R. McCormarch et al., 123–185. Baltimore: The Johns Hopkins University Press.

———. 1980. Anatomy of a Priority Conflict: The Case of Element 72. *Centaurus* 23: 275–301.

Krips, Henry. 1987. *The Metaphysics of Quantum Theory*. Oxford: Clarendon.

Kuhn, Thomas S. [1962]1969. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Lumsden, C., and E.O. Wilson. 1981. *Genes, Mind, and Culture*. Cambridge, MA: Harvard University Press.

Ladyman, James. 1998. What is Structural Realism. *Studies in History and Philosophy of Science* 29: 409–424.

Ladyman, James & Don Ross. 2007. *Every Thing Must Go. Metaphysics Naturalized.* With David Spurrett & John Collier. Oxford: Oxford University Press.

Laudan, Larry. 1982. A Confutation of Convergent Realism. *The Journal of Philosophy* 48, 19–48. Reprinted in J. Leplin (ed.) *Scientific Realism* (1984), 218–49. References are to this volume.

Lehmkuhl, Dennis. 2014. Why Einstein Did Not Believe That General Relativity Geometrizes Gravity. *Studies in History and Philosophy of Modern Physics* 46: 316–326.

Lewis, David. 1973. *Counterfactuals*. Cambridge, MA: Harvard University Press.

Lowe, E. Jonathan. 1998. *The Possibility of Metaphysics. Substance, Identity, and Time*. Oxford: Clarendon Press.

Lusanna, L. and M. Pauri. 2006. "Explaining Leibniz-equivalence as difference of non-inertial appearances: Dis-solution of the Hole Argument and physical individuation of point-events, *Studies in History and Philosophy of Modern Physics*, 37(4): 692–725.

Mackie, John L. 1972. *The Cement of the Universe*. Oxford: Clarendon Press.

Marks, Charles E. 1980. *Commissurotomy, Consciousness, and the Unity of the Mind*. Cambridge: MA, MIT Press.

Marsuzawa, T. 2007. Comparative Cognitive Development. *Developmental Science* 10: 97–103.

Maudlin, Tim. 1989. The Essence of Spacetime. In A. Fine and J. Leplin (eds.) *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association 1988*. Vol. 2, 82–91.

———. 1990. Substances and Spacetime. What Aristotle Would Have Said to Einstein. *In Studies in History and Philosophy of Science* 21: 531–561.

———. 2013. The Nature of the Quantum State. In *The Wave Function. Essays on the Metaphysics of Quantum Mechanics*, eds. Alyssa Ney and David Z. Albert. Oxford: Oxford University Press.

McMullin, E. 1984. A Case for Scientific Realism. In *Scientific Realism*, ed. J. Leplin, 8–40. Berkeley: University of California Press.

Melia, Joseph. 2000. Weaselling Away the Indispensability Argument. *Mind* 109: 455–480.

———. 2002. Response to Colyvan. *In Mind* 111: 75–79.

Michotte, Albert. [1945]1962. *The Perception of Causality*. Andover, MA: Methuen.

Mitchell, S. 1988. Constructive Empiricism and Anti-Realism. In A. Fine & J. Leplin, *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association 1988*, Vol. 1, 174–180.

Mukhanov, Viatscheslav. 2007. Cosmology and the many worlds interpretation of quantum mechanics. In *Universe or Multiverse?* ed. B. Carr, 267–274. Cambridge: Cambridge University Press.

Nagel, Thomas. 1974. What is it like to be a Bat? *Philosophical Review* 83: 435–450.

———. 1986. *The View From Nowhere*. New York: Oxford University Press.

Newman, M. 2004. Ramsey-Sentence Realism as an Answer to the Pessimistic Meta-Induction. In *Proceedings Philosophy of Science Assoc. 19th Biennial Meeting—PSA2004*: *PSA 2004 Contributed Papers*.

Newton, Isaac. [1729]1960. *Mathematical Principles of Natural Philosophy*. Berkeley: University of California Press.

———. 1962. *Unplublished Scientific Papers of Isaac Newton*, eds. Hall, A. R. and M.B. Hall. Cambridge: Cambridge University Press.

Newton-Smith, W.H. 1978. The Underdetermination of Theories by Data. *Aristotelian Society, suppl.* 52: 71–91.

———. 1981. In Defence of Truth. In *The Philosophy of Evolution*, (Series: *Harvester Studies in Philosophy*, vol. 26), eds. U.J. Jensen and R. Harré, 269–289. Bristol: The Harvester Press.

———. 1981. *The Rationality of Science*. London: Routledge and Kegan Paul.

Ney, Alyssa, and David Z. Albert. 2013. *The Wave Function. Essays on the Metaphysics of Quantum Mechanics*. Oxford: Oxford University Press.

Niiniluoto, I. 1977. On the Truthlikeness of Generalizations. In *Basic Problems in Methodology and Linguistics*, eds. R. Butts and J. Hintikka, 121–147. Dordrecht: Reidel.

Norton, John D. 1985. What was Einstein's Principle of Equivalence? *Studies in History and Philosophy of Science* 16: 203–246.

———. 1988. The Hole Argument. In A. Fine & J. Leplin (eds.) *PSA 1988* Vol. 2. 56–64.

———. 2004. The Hole Argument. In *Stanford Encyclopedia of Philosophy* http://plato.stanford.edu/

Perreault, Charles & Sarah Matthew. 2012. Dating the Origin of Language Using Phonemic Diversity. *Plus One*, April 27. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3338724/

Petersen, Andur S. 1985. Kan matematiske og fysiske teorier tolkes realistisk?. (Can mathematical and physical theories be interpreted realistically?). In *Nyere danske filosofi*, 1–19. Philosophia: Aarhus.

Pinker, Steven, and R. Jackendoff. 2005. The Faculty of Language: What's Special About It? *Cognition* 95: 201–236.

Pooley, Oliver. 2006. Points, Particles and Structural Realism. In *The structural foundations of quantum gravity*, eds. S. French, D. Rickles, and J. Saatsi, 83–120. Oxford: Oxford University Press.

Popper, Karl R. 1972. *Objective Knowledge: An Evolutionary Approach*. Oxford: Oxford University Press.

Provinelli, D.J., and J. Vonk. 2004. We Don't Need a Microscope to Explore the Chimpanzee's Mind. *Mind & Langauge* 19: 1–28.

Psillos, Stathis. 1995. Is Structural Realism the Best of Both Worlds? *Dialectica* 49: 15–46.

———. 1996. Scientific Realism and the 'Pessimistic Induction'. *Philosophy of Science* 63: S306–S314.

———. 2000. Carnap, the Ramsey-Sentence and Realistic Empiricism. *Erkenntnis* 52: 253–279.

Putnam, Hilary. 1978. *Meaning and the Moral Sciences*. London: Routledge and Kegan Paul.

———. 1981. *Reason, Truth and History*. Cambridge: Cambridge University Press.

———. [1981a]1983. Why Reason Cannot Be Naturalized. In his *Realism and Reason*. *Philosophical Papers.* Vol. 3, 205–228. Cambridge: Cambridge University Press.

———. [1981b]1983. Why Reason Cannot be Naturalized. In his *Realism and Reason*. *Philosophical Papers.* Vol. 3, 229–247. Cambridge: Cambridge University Press.

Quine, Willard Van Orman. [1951]1966. On Carnap's Views on Ontology. Reprinted in *The Ways of Paradox and Other Essays.* Cambridge, MA: Harvard University Press.

———. (1969). Existence and Quantification. In *Ontological Relativity and Other Essays,* 92–113. New York: Columbia University Press.

Ramsey, Frank P. ([1930]1990). *Philosophical Papers*. Edited by D. H. Mellor. Cambridge: Cambridge University Press.

Rickles, Dean. 2008. *Symmetry, Structure, and Spacetime*, (Series: Philosophy and Foundations of Physics. Vol 3). Amsterdam: Elsevier.

Rodriguez-Pereyra, G. 2006. Truthmakers. *Philosophical Compass* 1(2): 186–200.

Rohrlich, Fritz. 1986. Reality and Quantum Mechanics. *Annals of the New York Academy of Sciences* 480: 373–381.

Rovelli, C. 1997. Halfway through the Woods: Contemporary Research on Space and Time. In *The Cosmos of Science*, eds. J. Earman and J. Norton, 180–223. Pittsburgh: University of Pittsburgh Press.

Rubin, E., 1950, Visual Figures Apparently Incompatible with Geometry. *Acta Psychologica* 7: 365–387.

Ruse, Michael. [1986]1998. *Taking Darwin Seriously. A Naturalistic Approach to Philosophy*. New York: Prometheus Books.
———. 2012. *The Philosophy of Human Evolution*. Cambridge: Cambridge University Press.
Salmon, Wesley C. 1984. *Scientific Explanation and the Causal Structure of the World*. Princeton: Princeton University Press.
Searle, John. 1978. Literal meaning. *Erkenntnis* 13: 207–224.
———. 1995. *The Construction of Social Reality*. New York: The Free Press.
Sellars, Roy Wood. 1922. *Evolutionary Naturalism*. Chicago: Open Court.
Sellars, Wilfred. 1962. Philosophy and The Scientific Imagine of Man. In *Frontiers of Science and Philosophy*, ed. Robert Colodny, 35–78. Pittsburgh: University of Pittsburgh Press.Reprinted in *Science*, *Perception and Reality* (1963), 1–41.
Shettleworth, S.J. 2010. *Cognition, Evolution and Behaviour*. New York: Oxford University Press.
Sider, Theodore. 2011. *Writing the Book of the World*. Oxford: Oxford University Press.
Sklar, L. 1974. *Space, Time, and Spacetime*. Berkeley: University of California Press.
Smart, J.C.C. 1963. *Philosophy and Scientific Realism*. London: Routledge & Kegan Paul.
Sperry, R.W. 1970. Perception in the Absence of the Neocortical Commissures. *Perception and Its Disorders (Research Publication Association for Research in Nervous and Mental Disease)* 48: 123–138.
Strawson, P.F. [1950]1964. Truth. Reprinted in Pitcher, G. (ed.) *Truth*, Englewood Cliffs: Prentice-Hall.
Thagard, Paul. 1980. Against Evolutionary Epistemology *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association 1980*, Vol 1, 187–196.
Tegmark, Max. 2007. Many lives in many worlds. *Nature* 448: 23–24.
———. 2008. The Mathematical Universe. *Foundations of Physics* 38: 101–150.
Tipler, Frank. 2007. *The Physics of Christianity*. New York: Doubleday.
Tranekjær Rasmussen, E. 1955. On Perspectoid Distances. *Acta Psychologica* 11: 297–302.
Ulbæk, Ib. 1998. The Origin of Language and Cognition. In *Approaches to the Evolution of Language: Social and Cognitive Base*, eds. J.R. Hurford,

M. Studdert-Kennedy, and C. Knight, 30–43. Cambridge: Cambridge University Press.
van Fraassen, Bas C. 1980. *The Scientific Image*. Oxford: Oxford University Press.
———. 1985. Empiricism in the Philosophy of Science. In *Images of Science*, eds. P.M. Churchland and C.A. Hooker. Chicago: The Chicago University Press.
———. 1989. *Laws and Symmetry*. Oxford: Clarendon Press.
———. 1999. The Manifest Image and the Scientific Image. In *Einstein Meets Magritte: The White Book—An Interdisciplinary Refection*, ed. D. Aerts, 29–52. Dordrecht: Kluwer.
———. 2006. Structure: Its Shadow and Substance. *The British Journal for Philosophy of Science* 57: 275–307.
———. 2008. *Scientific Representation: Paradoxes of Perspective*. Oxford: Oxford University Press.
Vickers, Peter. 2012. Historical Magic in old Quantum Theory. *European Journal for Philosophy of Science* 2: 1–19.
Vlamings, P.H.J.M., J. Usher, and J. Call. 2006. How the Great Apes (Pan troglodytes, Pongo pygmaeus, Pan paniscus, and Gorilla gorilla) Perform on the Reversed Contingency Task: The Effects of Food Quantity and Food Visibility. *In Journal of Experimental Psychology: Animal Behaviour Processes* 32: 60–70.
Warnoch, G.J. [1962]1964. A Problem About Truth. Reprinted in G. Pitcher (ed.). *Truth*. Englewood Cliffs: Prentice-Hall, 54–67.
Williams, Donald C. 1953. On the Elements of Being I and II. In *Review of Metaphysics*, 17: 3–18 and 171–192.
Worrall, John. 1989. Structural Realism: The Best of Both Worlds? *Dialectica* 43: 99–124.
Wright, Crispin. 1987. *Realism, Meaning, and Truth*. Oxford: Basil Blackwell.
———. 1992. *Truth and Objectivity*. Cambridge, MA: Harvard University Press.
Zentall, Thomas R., Edward A. Wasserman, Olga F. Lazareva, Roger K.R. Thompson, and Mary Jo Rattermann. 2008. Concept Learning in Animals. *In Comparative Cognition and Behavior Reviews* 3: 13–45.

Index

Note: Page number followed by 'n' refers to notes.

J. Faye, *Experience and Beyond*,
DOI 10.1007/978-3-319-31077-0

Q

R