THEORIES AND REALITY:

FIVE ESSAYS ON QUINE AND UNDERDETERMINATION

BY

ROGÉRIO PASSOS SEVERO
B.A., Universidade Federal do Rio Grande do Sul, 1995
M.A., Universidade Federal do Rio Grande do Sul, 2000

DISSERTATION

Submitted as partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Philosophy
in the Graduate College of the
University of Illinois at Chicago, 2006

Chicago, Illinois

Para Bárbara, com amor

ACKNOWLEDGEMENTS

For comments, advice, and discussion, I am very much indebted to Peter Hylton, who served as my advisor, and to the other members of the dissertation committee: Bill Hart, Carl Hoefer, Daniel Sutherland, and Nick Huggett. I would also like to thank my colleagues Andrew Blom, David Harker, and Dearbhla Kelly, who were patient enough to discuss with me the topics of this dissertation at its early stages, and a number of other friends, teachers, and colleagues from whose insight and criticism I also benefited, including Alexander Dolnick, Alexandre Noronha Machado, Barbara Martin, César Schirmer dos Santos, Cameron Brewer, Carlos Briones, Charles Schubert, Craig Fox, David Potts, Gelson Roberto, Georgios Katsikatsos, Ken Martin, Kent Schmor, Márcio Teixeira, Maria Balcells, Mauro Engelmann, Neal Grossman, Paulo Faria, and Seishu Nishimura. The following pages contain what I regard to be a modest contribution to the studies of Quine and underdetermination. Those contributions would have been even slighter if it were not for the help I had from the people just mentioned.

RPS

TABLE OF CONTENTS

LIST OF ABBREVIATIONS OF WORKS BY W.V. QUINE

CAM	Commensurability and the alien mind
CB	Comment on Bergström
CGC	A comment on Grünbaum's claim
CLT	Carnap and logical truth
CNS	Comments on Newton-Smith
EC	Empirical content
EESW	On empirically equivalent systems of the world
EN	Epistemology naturalized
FLPV	From a Logical Point of View
FM	Facts of the matter
FME	Five milestones of empiricism
FSS	From Stimulus to Science
GWW	Goodman's Ways of Worldmaking
ITA	Indeterminacy of translation again
LC	Lectures on Carnap
NLWM	Naturalism; or, living within one's means
NNK	The nature of natural knowledge
POS	In praise of observation sentences
PR	Posits and reality
PT	Pursuit of Truth
PTF	Progress on two fronts
R	Responses
RA	Relativism and absolutism
RC	Reply to Chomsky
RIT	On the reasons for indeterminacy of translation
RJJCS	Reply to J.J.C. Smart
RJV	Reply to Jules Vuillemin
RJW	Reply to John Woods
RPR	Reply to Paul A. Roth
RR	The Roots of Reference

LIST OF ABBREVIATIONS OF WORKS BY W.V. QUINE (continued)

RRG	Reply to Roger F. Gibson, Jr
SLS	The scope and language of science
SN	Structure and nature
TDE	Two dogmas of empiricism
TDR	'Two Dogmas' in retrospect
TI	Three indeterminacies
TT	Theories and Things
TTPT	Things and their place in theories
UPM	Use and its place in meaning
VITD	On the very idea of a third dogma
WB	Quine, W.V., Ullian, J.: The Web of Belief
WO	Word and Object
WPB	What price bivalence?
WPO	Whither physical objects?
WTI	On what there is

SUMMARY

The thesis of underdetermination has received considerable attention in the philosophical literature of the last few decades, yet little consensus has been reached about how it is to be thought of or formulated. Even its most influential proponent, W.V. Quine, changed his mind several times about how to characterize the thesis, and vacillated about some of its implications. The present dissertation contains a comparative analysis of Quine's views on underdetermination and contemporary debates on the matter.

Five main points are defended in the dissertation: (i) holism and underdetermination are distinct theses; while the former may be thought to lend credence to the latter, it is insufficient to establish it; (ii) Quine's formulation of the thesis is significantly weaker than commentators have thought; (iii) rivalry among empirically equivalent theories is best understood in terms of non-intertranslatability of fundamental terms or predicates; (iv) some recent criticisms of the thesis are empty in that they unduly strengthen the thesis so as to render it untenable; and some recent defenses of the thesis rely on an overly lax view of what counts as a rival theory, or unduly weaken it in such a way that it ends up indistinguishable from radical skepticism; (v) contrary to a common assumption in the recent literature, underdetermination does not have to be thought as incompatible with epistemic realism.

1. INTRODUCTION

About the thesis of underdetermination one could perhaps say what W.D. Hart once wrote of the Löwenheim-Skolem theorem: "rather a philosophical tease; it has always excited philosophical imaginations, but without ever quite delivering on its suggestive intimations" (1970, p. 98). The thesis itself has been variously formulated. Very roughly it says that no matter how well supported by observations a theory of the world may be, alternative theories exist that are equally warranted by observations. Observations alone do not determine one theory above all possible alternatives; theories are *underdetermined* by observations. At first glance, the thesis seems plausible, perhaps even undeniable. Surely alternative explanations for the same observations are possible, even likely. And it is plausible to think that unknown alternative theories exist that remain forever unborn.

Over the last half-century, the philosophical literature has registered a rather intense debate about the meaning of the thesis. Despite its seeming plausibility, philosophers have found numerous difficulties both in giving it a precise formulation and even more so in finding for it an adequate justification.[1] Additionally, the thesis has raised a number of concerns because of its apparent implications. One major concern has been whether the thesis, if true, can be reconciled with a scientific realism, the doctrine or view commonly accepted by practicing scientists that the theoretical entities and principles posited by science are real. If alternative theories are equally warranted by observations, then it would seem that the certainty with which those entities and principles are posited is unwarranted; perhaps only what we observe can be ascertained as real, while all else could be just fabrications of our collective imaginations.

[1] An example is Quine's EESW. See also Humphries (1970), Bar-On (1986), Ben-Menahem (1990), Laudan and Leplin (1991), Earman (1993), Hoefer and Rosenberg (1994), Norton (1994), Stanford (2001), and Magnus (2003).

Accordingly, some philosophers have argued that if the thesis of underdetermination holds, then we must back down on one aspect or another of scientific realism. Our knowledge claims about the world would have to remain limited to what we can directly observe. The theoretical entities and principles posited by science, it is argued, should not be taken to be existent and real, but as mere tools for the reliable prediction of observations. Assertions about those theoretical principles and entities may be false or misleading. Further claims about reality or the structure of reality or whatever lies beyond the reach of direct observation would be speculative at best.[2]

Other philosophers have opted instead to defend scientific realism *against* underdetermination, by systematically pointing out problems with the intelligibility and justification of the latter.[3] The success of science in predicting observations and the lack of relevant or viable alternatives that came to be held in modern and contemporary times are presented as evidence in favor of the warrantedness of theories that rival our own. Moreover, because most of the predictions of observations in contemporary science can only be made by making use of some unobservable entities, the success of science and the lack of relevant or viable alternative theories is also seen as evidence for our capacity to know an independently existing world even beyond that which we directly observe. The thesis of underdetermination seems to undermine these 'realistic' views of science. Hence, some realist philosophers of science have repeatedly questioned the claim that rival theories can be empirically equivalent. They argue that although it may be true that observations do not determine one theory over all possible alternatives, at each moment of our history only a few alternative theories have actually

---

[2] See van Fraassen (1980), Bergström (1990), Kukla (1998), Stanford (2001), and Okasha (2002).

[3] See Newton-Smith (1978), Laudan and Leplin (1991), Kitcher (1993, chapter 7.6), Norton (1993, 1994, 2003c), and Massimi (2004).

been presented as rivals. For those relevant cases, observations coupled with theoretical constraints and methodological considerations about theory construction – such as simplicity, generality, modesty, and fecundity – have always seemed to favor more or less clearly one theory above its rivals.[4] Another argument points out that the empirical content of a theory varies according to the technology available for gathering observation reports. Hence, whether two theories are empirically equivalent also varies according to the state of science and technology in general. The claim that two rival theories are empirically equivalent and therefore underdetermined by observations would thus always be defeated by future technological developments.[5] Underdetermination might at best be a transient phenomenon.[6]

So in the literature we find some authors defending underdetermination and claiming that it carries with it some substantial philosophical consequences, such as anti-realism or agnosticism about theoretical entities and principles. We find also authors claiming that underdetermination is systematically defeated by the maxims and constraints of theory construction. Despite those conflicting attitudes, both those views abide by a common presupposition, namely, that underdetermination is incompatible with scientific realism. Given that presupposition, it is not surprising that the main questions tackled in the recent literature on the topic are concerned with what to give up or adjust: underdetermination or scientific realism.

In some regards, the thesis of underdetermination can be compared to Benacerraf's dilemma. Mathematical objects are abstract, and therefore knowledge of them cannot be acquired through what are taken to be the usual sensory means. Hence, it would seem that we

---

[4] See Norton (1993 and 2003c), Cordero (2001), and Massimi (2004) for discussion and examples.

[5] See Laudan and Leplin (1991), Leplin (1997a), and Massimi (2004).

[6] The notion of transient underdetermination can be found in Sklar (1981) and Stanford (2001).

face the following dilemma about the objects of mathematics about the objects of mathematics: either we do not take them to be objects in the same sense as we take empirical objects to be objects, or we do take them as objects in the ordinary sense but give up the idea that they can be known by sensory means – or known at all, if one restricts knowledge to what we can acquire by sensory means. Benacerraf's dilemma thus opposes a realist belief about the ontological status of mathematical objects to an empiricist epistemology of those objects.[7] Some have held that in mathematics one must either choose to be an empiricist in epistemology and give up realism in metaphysics or choose realism in metaphysics and give up empiricism in epistemology.[8] Underdetermination has been thought to pose a similar dilemma. If more than one theory can be equally warranted by observations, then it seems that either one should accept the empiricist epistemology which yields that result and give up the idea that the theoretical entities posited by science are real – or at least that we can know them – or one should accept that those objects can be known but must reject the claim that the theories which posit them are underdetermined. Like Benacerraf's dilemma in regards to mathematics, underdetermination would seem to reveal an inconsistency between empiricism and realism within natural science.[9]

Rather at odds with the prevailing views on underdetermination is the work of W.V. Quine. He was the most important and influential proponent of the thesis in the twentieth century, and one of the first to analyze it in some detail. Quine was also one of the first to identify some of the shortcomings which have motivated the kind of opposition to the thesis

---

[7] 'Empiricist epistemology' here just means a theory of knowledge which takes sensory input to be the source of our knowledge of the world. 'Empiricism', here, should not be taken to be opposed to realism by definition, but as imposing constraints on knowledge which may be incompatible with a realist attitude towards unobservable and abstract objects.

[8] For further discussion, see Benacerraf (1973) and Hart (1991).

[9] Views along these lines are expressed by Gibson (1980 and 1986) and Bergström (1984).

found in the recent literature. As with other Quinean doctrines, however, his views on underdetermination have been often misunderstood. This is in part due to the systematic nature of Quine's philosophy: Specific theses and doctrines which he holds are not easily detachable or even distinguishable from other theses and doctrines within his philosophy. Misunderstanding of Quine on underdetermination seems to have more specific reasons as well. One reason is that over the years he changed his formulation of the thesis at least twice, and vacillated about some of its implications on various occasions. This has rendered his views on underdetermination somewhat opaque and hard to comprehend. A second, perhaps more important reason, is that his views on the matter diverge from most that have been registered in the recent literature. In particular, Quine does not see the thesis as entailing the falsehood of scientific realism. In fact, for Quine the thesis does not seem to have any significant metaphysical or epistemological implications. The thesis itself, he argued, is untenable unless it is formulated in terms that make it weaker than most versions currently debated. Some stronger versions of the thesis currently held in the literature are for Quine no more than conjectures. The version he did maintain was only a "vague and modest" (EESW, p. 327) one, as we shall see (chapters 2 and 3). Nevertheless, Quine thought it was a plausible thesis, given what we know about the "less-than-rigid" (PR, p. 254) connections that obtain between theories and observations. Those views on underdetermination did not prevent Quine from describing himself as a realist, indeed a "robust" realist (TTPT, p. 21). Underdetermination and realism were not doctrines Quine regarded as incompatible.

The present dissertation is an analysis of the two sets of topics mentioned above: on the one hand, the current literature on underdetermination; on the other, Quine's views on the matter. Debates on the interpretation of Quine's philosophy and debates on the nature and implications

of underdetermination have been taking place in relative isolation from one another. One of the goals of this dissertation is to show how a comparative analysis of contemporary debates on the meaning and implications of the thesis and of Quine's views may benefit both sides. Some of Quine's views on the matter were left undeveloped, and the recent literature contains a number of suggestions that can easily complement them. Also, some aspects of Quine's views need further clarification and adjustments, some of which are already available in the literature. On the other hand, the philosophical debate concerning the nature of underdetermination can benefit much from Quine's suggestions and insights. There is still confusion about what exactly underdetermination is and how best to formulate the thesis, and these are certainly issues about which Quine has something to say.

The dissertation contains five more or less self-contained chapters. Each may be read separately from the others, and has a point of its own. Collectively, the five chapters contain a defense of a modest, Quinean, formulation of the thesis, and an analysis of the thesis which shows that from a naturalistic standpoint realism and underdetermination are not incompatible. Additionally, some arguments against the thesis of underdetermination, and some others from underdetermination to antirealism, which are common in the current literature, are rejected in this dissertation. The contributions that each of the five chapters makes to the current debates on Quine and underdetermination can be briefly summarized as follows:

Chapter 2: In TDE Quine wrote that holism is his "countersuggestion" to the dogma of reductionism as it figures in "Carnap's doctrine of the physical world in the *Aufbau*" (p. 41). Holism is thus present in Quine's philosophy at least since the early 1950s, and was maintained

throughout thereafter. It is a fundamental doctrine within his system both in the sense that most other doctrines within his philosophy are either dependent on it or are deeply affected by it, and in the sense that Quine took it to be obviously true. In other works (PR, *WO*, NNK, EESW), Quine then goes on and argues that holism "lends credence" to the thesis of underdetermination. He also shows, however, that holism is insufficient to establish the truth or even the intelligibility of underdetermination. Quine is widely misread on this particular point, and taken to have asserted that underdetermination is entailed by holism. The thesis of holism can of course be formulated in various ways. The version which Quine defended is relatively weak. Still, even a stronger version of holism is insufficient to entail underdetermination. Chapter 2 addresses these issues and attempts to clarify Quine's views on holism and underdetermination, distinguishing the two and correcting some common misinterpretations. Chapter 2 also discusses – although only very briefly – Quine's initial formulations of holism in connection with his reading of Carnap's *Aufbau*, and how his thesis of holism differs from that of Duhem.

Chapter 3: Quine revised his formulation of the thesis of underdetermination at least twice and at least three times he changed his mind about what to say of the truth of empirically equivalent rival theories. Those changes occurred over a period forty years or so, and they only stopped happening when Quine stopped writing in the 1990s. Only very recently have commentators been able to assess Quine's views on underdetermination comprehensively, and fully analyze the changes they underwent. A few authors have partially tracked down those changes;[10] chapter 3 complements the work of those authors in the light of Quine's writings published in the late 1980s and 1990s. Additionally, it is argued that Quine's views on underdetermination are in general weaker than they are usually taken to be in the philosophy of

[10] Most notably, Gibson (1988 and 1998) and Bergström (1990 and 1993).

science literature. One reason for that somewhat diffused misrepresentation of Quine's views on the matter lies precisely in the fact that they underwent changes which were not later systematically laid out or explained by Quine himself. Another reason lies in the fact that Quine wrote about two different sets of issues at different times. Up until 1975, when EESW was published, Quine's writings on underdetermination dealt mostly with the formulation and justification of the thesis. Afterwards, most of his writings concern a different issue: *given underdetermination*, how are we to assess the truth of empirically equivalent rival theories? The latter question was one about which he vacillated between what he called an 'ecumenical' attitude, which assigns truth to any theory that is empirically equivalent to our own, and a 'sectarian' attitude, which assigns truth only to whichever theory we hold at the moment. In the secondary literature, however, Quine's writings on the truth of rival theories are often read as addressing the formulation and justification of the thesis as well. As a result, he is frequently thought to have put forth a much stronger thesis of underdetermination than the one he actually did. In chapter 3 it is argued that Quine maintained a mitigated ("vague and modest") version of the thesis at least since 1975, and that his vacillations on the truth of empirically equivalent rival theories should be interpreted accordingly. His writings on the question of the truth of rival systems of the world have no bearing on the formulation and justification of the thesis.

Chapter 4: The hypothesis that no matter how warranted by observations a global theory is, it admits of empirically equivalent rivals, was put forth by Quine as a plausible conjecture. The meaningfulness of that conjecture, however, and in particular of the notions of 'empirical equivalence' and 'rivalry' which its formulation employs, is not clear on a first reading. In fact, some authors have claimed that the thesis must remain defeasible by future technological developments, and that the notion of 'empirical equivalence' employed by Quine

and others is confused, the result of an oversimplification.[11] Other authors have argued that the thesis of underdetermination is in principle false, because any two empirically equivalent theories, there would be a way of rendering them logically equivalent through translation.[12] Empirical equivalence would then entail logical equivalence. The very idea of empirically equivalent yet *rival* theories is thus challenged: empirically equivalent rival theories would be at most terminologically distinct, but nonetheless intertranslatable versions of the same theory. Quine was aware of the difficulty, and mitigated his thesis of underdetermination accordingly (EESW, pp. 318-327). Still, he maintained that there are (in a platonic sense, so to speak) empirically equivalent rival theories of the world which, if we were to find them, we would be unable to render logically equivalent through translation. Thus conceived, the thesis of underdetermination becomes vulnerable to a criticism originally targeted at the so-called 'thesis of incommensurability', defended by Kuhn (1970 [1962]) and Feyerabend (1962). That thesis says that competing theories may contain terms and phrases which cannot be translated into the language of their rivals. Incommensurability was challenged by Putnam (1975 and 1981) and Davidson (1984), who argued that the notion is incoherent, especially when rendered in terms of untranslatability as Kuhn eventually did (2000a). Chapter 4 reformulates some of the responses by Kuhn and Feyerabend to criticisms of the thesis incommensurability so as to address the criticisms directed at the thesis of underdetermination as well. Those reformulated responses further the understanding of underdetermination and empirical equivalence against authors who have recently dismissed those notions as irrelevant and philosophically inane. Additionally, some illustrations of the thesis of incommensurability are presented in chapter 4 as evidence for

---

[11] For example, Laudan and Leplin (1991). See also Kitcher (1993, pp. 247-256), Stanford (2001), Magnus (2003), Norton (2003c), and Massimi (2004).

[12] See Dummett (1981, p. 618 n).

the thesis of underdetermination, or at least of its intelligibility. The chapter thus complements Quine's defense of the thesis, and provides additional insight into what the thesis can be thought to entail.

Chapter 5: In the recent literature, two attitudes towards underdetermination stand out: (i) underdetermination is false, and (ii) underdetermination entails anti-realism or about theoretical entities and principles. A closer look at the published articles defending each of those views reveals that the less sympathetic an author tends to be toward the thesis, the stronger her formulations of the thesis of underdetermination will be; conversely, the more sympathetic the author, the weaker the constraints she imposes on the thesis. As a result, the rather pervasive disagreements about the truth of the thesis which we find in the literature seem to be fueled by a lack of common understanding about how to formulate the thesis. Chapter 5 attempts to sort out the various formulations of the thesis currently found in the literature, and show that some of the criticisms directed at stronger versions of the thesis are empty: they are criticisms of an untenable thesis defended by no one. It is also argued that some of the weaker versions of the thesis defended in the literature are either too weak to be called "thesis of underdetermination" or rely on implausibly lax views on what counts as a rival theory. Chapter 5 discusses shortcomings in the justification of several versions of the thesis, and shows that Quine's mitigated formulation remains the most plausible and tenable version of the thesis available in the literature.

Chapter 6: As mentioned above, in the current literature, the thesis of underdetermination is generally perceived to be inconsistent with scientific realism. Scientific realists tend to say that there are good reasons to construe the unobservable parts of a model as true, if the theories that posit them are thoroughly confirmed by observations. Oftentimes,

however, underdetermination is formulated in such a way that it entails the possibility of logically *incompatible* theories, each positing its own set of theoretical entities and principles. Given that formulation, it would seem that observations can warrant theories that are logically incompatible and that one is therefore forced either to suspend judgment about the truth of rival, empirically equivalent theories, or to take them as mere instruments of prediction of observations. In either case, one would thereby give up part of what scientific realism is thought to entail. Chapter 6 explores a Quinean response to the problem, and defends it against some recently published criticisms. Quine argued that the notion of logical incompatibility, when applied to empirically equivalent global theories, proves inconsequential. Divergencies among empirically equivalent theories can only occur at the level of theoretical statements, which hinge on observations only indirectly. But then logical incompatibilities can be altogether avoided by treating all pairs of incompatible sentences as cases of equivocation. Both the equivocation and the incompatibility can be avoided by adjusting the vocabulary of one of the theories. The conflict that may then remain between the two will at most be one of non-intertranslatability: two theories, T and T', may be such that no known manual of translation can render all the sentences of one logically equivalent to a sentence or conjunction of sentences of the other. Quine himself did not rule out the possibility that in principle one could always later find more sophisticated manuals of translation which render the two theories logically equivalent. The conflict which non-intertranslatable but empirically equivalent theories may be thought to have might therefore be merely apparent, and due exclusively to our ignorance of more sophisticated manuals of translation. In the absence of such manuals, however, two theories may legitimately, even if momentarily, be regarded as rivals. Each will assert some theoretical sentences which will have no known counterpart in the technical vocabulary of the rivals. Sentences which are true in one

theory cannot even be stated within the rival theory. The conflict, in such cases, is not one of logical incompatibility, but of alternative vocabularies which cannot be reduced one to another.

In the absence of logical incompatibility, theoretical conflicts among empirically equivalent theories appear devoid of the more substantial problems often thought to be associated with the thesis of underdetermination. Coexisting, alternative taxonomies and linguistic structures have been a common trait of the history of our cultures anyway. One does not thereby conclude that alternative worlds coexist. Likewise, empirically equivalent rival theories can be understood as an extreme or limiting case of linguistic variations. As we shall see, this suggests that the perceived conflict between underdetermination and scientific realism may simmer down, "bathetically", as Quine says (*PT*, p. 101), to a question of words. This is not to say that no problems or difficulties remain for a Quinean treatment of underdetermination and realism. To be sure, some are discussed at the end of chapter 6. However, those problems do not support some of the stronger criticisms that have been directed at Quine's views.[13]

In connection with this last point, chapter 6 also discusses Quine's description of himself as a "robust" realist. This has seemed puzzling to some,[14] so a brief explanation of what Quine meant is in order. Additionally, it is shown why some of the arguments from underdetermination to anti-realism and agnosticism currently offered in the literature are untenable from the perspective of a naturalist philosophy such as Quine's. Those arguments employ a notion of reality which is often too much disconnected from the standards of evidence within which it emerges and acquires its utility.

---

[13] For example, by Gibson (1986), Bergström (1984 and 1993), and Yalçin (2001).

[14] Sarkar (2000, p. 189), for example, suggests that Quine is an anti-realist.

The dissertation concludes with a brief discussion of the relevance and philosophical significance that the thesis of underdetermination can be thought to have. Some overlap among the five chapters has been tolerated, so as to preserve the more or less self-contained nature of each one of them.

2. HOLISM AND UNDERDETERMINATION

> "[T]he point of philosophy is to start with something so simple as not to seem worth stating, and to end with something so paradoxical that no one will believe it."
>
> (Russell, The Philosophy of Logical Atomism, p. 53)

2.1 **Introduction**

Both the thesis of holism and the thesis of underdetermination have been very much associated with the philosophy of Quine. Indeed, he was the most important and influential proponent of both in the twentieth century.[15] Quine presented holism as an alternative to the dogma of reductionism, the view that the empirical content of each individual cognitive statement can be reduced to statements containing only terms for sense-data and logical terms. More generally, holism is an alternative to the view – sometimes dubbed "atomism" – that each scientific statement is endowed with an empirical content of its own.[16] Against atomism, holism says instead that in general empirical content attaches not to individual sentences or terms but to more or less large sets of sentences, the whole of science or "chunks" of it (TDR, p. 268).[17] Throughout Quine's work, holism is presented as a trivial doctrine, undeniable, and obviously true. In response to an objection by Grünbaum, for example, Quine wrote:

---

[15] Prior to Quine, Duhem (1954 [1914]) and others – for example, Carnap (2002 [1937], p. 318) – had defended holism, and for that reason it is also known as the "Duhem-Quine thesis". Duhem's version of the thesis was narrower in scope than Quine's, however, and the philosophical use he made of it was more restricted (see below, sections 2.3 and 2.4). Some authors distinguish "confirmation holism" and "meaning holism": see, e.g., De Rosa and Lepore (2004, p. 66). In what follows, we shall be concerned primarily, if not exclusively, with what those authors call "confirmation holism".

[16] See Fodor and Lepore (1992, chapter 2), and Hylton (2002, pp. 12-13).

[17] In a well-known passage in TDE, Quine wrote that "our statements about the external world face the tribunal of sense experience not individually but only as a corporate body" (p. 41). The suggestion was then that empirical content attaches only to the whole of science. In later works, however, the thesis is mitigated somewhat. In TDR, for example, he refers to those statements as a "needlessly strong" (p. 268) assertion of holism, and concedes that empirical content can attach also to portions and branches of science, and to individual observation sentences and observation categoricals. See also EN, p. 79.

> For my own part I would say that the thesis [of holism] as I have used it *is* probably trivial. I haven't advanced it as an interesting thesis as such. I bring it in only in the course of arguing against such notions as that the empirical content of sentences can in general be sorted out distributively, sentence by sentence, or that the understanding of a term can be segregated from collateral information regarding the object. (CGC, p. 132)

In later writings, the obviousness of the doctrine is reaffirmed; in *PT*, for example, Quine wrote that "[i]t is difficult to see how anyone can question holism, in the sense now before us" (p. 16).[18] Despite its alleged obviousness, holism is a fundamental doctrine within Quine's system. Most aspects of his philosophy are directly affected by it: Holism is straightforwardly connected to Quine's qualms about Carnap's use of the analytic/synthetic distinction, to the theses of indeterminacy of translation and ontological relativity, and it has implications that directly affect Quine's naturalized epistemology, and his overall conception of philosophy as continuous with science.[19]

By characterizing holism as trivially and obviously true, and then drawing numerous implications from it, Quine's use of the thesis illustrates well the kind of activity which characterizes philosophy in the passage by Russell quoted above. Accordingly, some of the alleged implications of holism have seemed unbelievable to Quine's readers. Of those implications, perhaps the most polemical is the claim that scientific theories are underdetermined by observations, the so-called thesis of *underdetermination*. Quine himself did not maintain that holism *entails* underdetermination, but he did claim that holism renders it plausible (EESW, p. 313), and thus provides a good reason for asserting it. However, unlike holism, the thesis of underdetermination is not trivially true, and both its formulation and its justification are rather

---

[18] See also EESW (pp. 313-314).

[19] These connections are made explicit by Quine himself in TDE, EN, and NLWM, for example, and also *WO* and *PT*. For discussion, see Gibson (1988, 1996, and 1998) and Hylton (1982 and 2002).

controversial, as the recent literature can attest.[20] Even within Quine's philosophy, underdetermination is neither trivially true nor a fundamental doctrine, and it does not affect other aspects of his philosophy in the same pervasive way as holism does.

Something similar to what we see within Quine's system can also be seen in the philosophy of science literature more generally: modest versions of the thesis of holism are more or less widely accepted, while the thesis of underdetermination is often seen as problematic, untenable even. There are discrepancies in the ways the thesis of underdetermination has been formulated in the literature, and divergencies about what its implications are and whether they obtain. The thesis of holism is nonetheless generally perceived as suggesting some version or other of the thesis of underdetermination, at least *prima facie*, even by those who then go on to reject underdetermination. Some of the authors who reject underdetermination have also come to question holism itself, especially if the two theses are perceived as entailing some sort of uncertainty about the results of science.[21] Oftentimes those criticisms are facilitated by characterizations of the two theses which systematically blur their differences. An example of that kind of blurring can be found in Laudan (1990), who defines "Quinean underdetermination" as the thesis that "[a]ny theory can reconciled with any recalcitrant evidence by making suitable adjustments in our other assumptions about nature" (p. 274). As we shall see below, at best this is a statement with which Quine would reluctantly agree, but only as an imprecise characterization of *holism*, not *underdetermination* as Laudan suggests.[22] Underdetermination is not a thesis about the reconciliation of hypotheses or theories with observations, but about the

---

[20] See, for example, Laudan and Leplin (1991), Norton (1993 and 1994), Hoefer and Rosenberg (1994), Kukla (1998), Stanford (2001), Okasha (2002), and Magnus (2003).

[21] An example is Norton (1994).

[22] Similar confusions can be found in Kitcher (1993, p. 251), Leplin (1997, p. 210), Sarkar (2000, p. 187), and Norton (2003c).

possibility of alternative theories that predict the same observations. Strictly speaking, the thesis stated by Laudan is false even as a characterization of holism, since some theories clearly imply false observation categoricals, and are therefore irreconcilable with observations. Quine's thesis of holism says rather that individual hypotheses or non-observation sentences – but not whole theories, or at least not global theories – can always be reconciled with observations by making adjustments in other sentences of the original theory, or in the background of accepted theory and ordinary assumptions. Adjustments to a theory as a whole yield new theories, not the same theory; so the claim cannot simply be that any theory can be reconciled with any evidence.

The main goal of this chapter is to formulate both the thesis of holism and the thesis of underdetermination as clearly as possible, sorting out the connections and differences between the two. In so doing, we shall be arguing against authors who have blurred the distinction between them, and laying out the ground for a defense of moderate versions of both. Section 2.2 introduces the thesis of holism, focusing primarily on Quine's formulation; section 2.3 briefly recalls the origins of the thesis of holism in Duhem's work, and in Quine's reaction to Carnap's project in the *Aufbau*; section 2.4 formulates the thesis of underdetermination and explains why holism renders it plausible but is insufficient to establish it; and section 2.5 outlines some of the main consequences of each thesis.

2.2 **Holism**

Both holism and underdetermination have been variously formulated. Perhaps it would be more appropriate to speak of *theses* of holism and *theses* of underdetermination instead, since the formulations of each are not all equivalent. Very broadly speaking, however, holism may be understood as the thesis that evidence bears primarily on theories as wholes, and not on the

individual sentences of a theory. Sometimes this point is made by saying that the notion of 'empirical content' attaches in general not to individual sentences but to more or less large sets of sentences. Of course, this latter characterization turns on the notion of 'empirical content', needs to be clarified.

In the debates on the topic, the *content* of a sentence or set of sentences is usually taken to be the set of sentences which it implies that are neither logically valid nor analytic.[23] Logically valid and analytic sentences are thought to follow from any other sentence or set of sentences, regardless of what those say. In determining the content of "there are zebras in Africa", for example, it is irrelevant that "it is raining or not raining", or that "no bachelors are married" logically follow from it, but it *is* relevant that some synthetic sentences also follow from it. For example, "there are zebras in Africa" implies, given what we take zebras to be, that "there are mammals in Africa" and "there are vertebrates in Africa". The content of a sentence can thus be defined as the set of synthetic sentences that it implies. This definition can be trivially extended to sets of sentences: the set of synthetic sentences that follow from a given set of sentences.

The notion of 'content' can then be used to define the *empirical* content of a sentence or set of sentences: not the set of synthetic sentences that are implied by a given sentence or set of sentences, but some favored subset of such sentences held to be more directly confirmed or refuted by observations. Naturally, there is much room here for divergencies about which sentences are actually confirmed or refuted by observations, and even about what counts as an

---

[23] See, for example, Carnap (2002, § 49) and Sellars (1980, p. 266). Naturally, there are various ways of characterizing the notion of 'analyticity'. The issue is not particularly relevant for our purposes here. An historical analysis can be found in Coffa (1991); Quine's views can be found, e.g., in TDE, CLT, and *PT* (pp. 54-56).

observation. It is not clear whether any precise borderline can be conceived, even in principle. In an attempt to circumvent those difficulties, Quine avoids talking of observations and of sentences being confirmed (or refuted) by observations. He settles instead for the notions of 'observation sentences' and 'observation categoricals'. Observation sentences are sentences like "It's raining", "It's getting cold", and "That's a rabbit". They are *occasion sentences*, true on some occasions and false on others, and

> … should command the subject's assent or dissent outright, on the occasion of a stimulation in the appropriate range, without further investigation and independently of what he may have been engaged in at the time. A further requirement is intersubjectivity: unlike the report of a feeling, the sentence must command the same verdict from all linguistically competent witnesses of the occasion. (*PT*, p. 3)

What Quine calls "observation categoricals" are general sentences of made up of observation sentences.[24] These are sentences of the form "Whenever this, that": "Whenever there is smoke, there is fire", for example. The "whenever", Quine notes,

> … is not intended to reify times and quantify over them. What is intended is an irreducible generality prior to any objective reference. It is a generality to the effect that the circumstances described in one observation sentence are invariably accompanied by those described in the other. (*PT*, p. 10)

As such, Quine argues, observation categoricals can be learned very early on, even by a child before she has fully acquired something like quantification theory. They are unlike observation sentences in that they are not occasion sentences, but *standing sentences*, true or false once and for all. They can be immediately rejected if some occurrence of the first occasion is not followed by an occurrence of the second; and presumably command assent if the relevant occurrences systematically follow one another in the right order. As such, they play a crucial role in the testing of scientific theories. Scientific theories typically do not imply observation

---

[24] In works published in the 1970s and before, Quine was using the "less fruitful" notion of 'observation conditional', which is made up not of observation sentences, but of "standing sentences". For further clarification and discussion, see *EC*, pp. 26-27.

sentences, since theories in general comprise general statements or laws only, which are standing sentences, and therefore do not entail statements about particular occasions but only general relations among occasion types. Scientific theories do, however, imply observation categoricals, and it is by way of these that they come to be tested. Such, at least, is the way Quine broadly conceives the relations between theories and observations, while granting that this is an idealization which highlights only the most general features of that relation (see EESW, pp. 314 ff.; EC, *passim*; and TDR, p. 268).

Given these characterizations of content, observation sentences, and observation categoricals, Quine defines the notion of 'empirical content' rather straightforwardly as follows:

> Call an observation sentence *analytic* for a given speaker if, as in 'Robins are birds', the affirmative stimulus meaning for him of the one component is included in that of the other. Otherwise *synthetic*. Call a sentence or set of sentences *testable* if it implies some synthetic observation categoricals. Call two observation categoricals *synonymous* if their respective components have the same stimulus meanings. Then the *empirical content* of a testable sentence or set of sentences for that speaker is the set of all the synthetic observation categoricals that it implies, plus all synonymous ones. I add the synonymous ones so that merely verbal variation will not obstruct sameness of content. (*PT*, pp. 16-17)

Quine offers no definition of empirical content for sentences that are not testable, or for theories that do not imply testable observation categoricals because of vagueness in the formulation. Yet he nonetheless uses a related, more loosely defined notion of empirical content in contexts where he discusses more abstractly the notion of empirical equivalence. Sentences and sets of sentences can be said to be empirically equivalent for an individual if they imply the same set of synthetic observation categoricals, and those sentences and sets of sentences will be empirically equivalent "for a whole community when equivalent for each member" (*PT*, p. 17). This latter notion, of 'empirical equivalence', is one that Quine uses to speak more abstractly of theories, regardless of questions about the testability of theories. Theories can be said to be

empirically equivalent even if it is not quite clear how to test them. It suffices that the relevant scientific community agrees that evidence for or against one theory counts also as evidence for or against another:

> … much solid experimental science fails of testability in the defined sense. This can happen (…) because of vague and uncalibrated probabilities in the backlog of theory. No doubt it happens also in more complex ways, not clearly understood. I have no definition of empirical content to offer for such theories, but it still seems to make reasonable intuitive sense to speak of empirical equivalence among them, since experimentation is still brought to bear. The idea is that whatever observation would be counted for or against the one theory counts equally for or against the other. (*PT*, pp. 95-96)[25]

Given Quine's definition of empirical content, the thesis of holism can now be stated more precisely: *in general* scientific, or cognitive, sentences lack empirical content when considered individually, or in isolation from other sentences. In general they do not, on their own, imply observation sentences or synthetic observation categoricals. The "in general" above is needed to accommodate two exceptions. The first is that some scientific statements are themselves observation sentences or synthetic observation categoricals. Since they imply themselves, they do have an empirical content of their own even when considered individually. The second exception is that individual sentences that lack empirical content may be conjoined into longer sentences that do have empirical content. One may even conjoin all the sentences that make up a scientific theory, or large portions of them, into a very long conjunctive sentence. That sentence then should have an empirical content of its own.[26]

Granted those two exceptions, holism has an almost trivial justification. Most sentences that make up the corpus of a scientific theory can only imply observation sentences and synthetic

---

[25] This extended use of the notion of 'empirical equivalence' appears in Quine's formulation of the thesis of underdetermination; some of the difficulties associated with that thesis turn on the extended use of the notion. More on this issue below (section 2.4).

[26] See RJV, p. 620.

observation categoricals when taken together with a number of other sentences and ordinary assumptions. Take, for example, the statement that most living plants produce oxygen in the presence of sunlight. This is true of most living plants, but false of some plants that are parasitic on other plants, and of some plants which extract their nutrients from dead organic matter. The general claim on its own, however, implies no observation sentences or synthetic observation categoricals, since it says nothing about what oxygen is and how one detects it or what living plants are and how they are identified. In the absence of such specifications, the claim that most living plants produce oxygen in the presence of sunlight fails to imply any observation sentence or synthetic observation categorical. A very large number of sentences is needed for that implication to obtain, including sentences specifying what oxygen and plants are. Only together do those sentences have some empirical content. But once those sentences are added, what one has is a whole theory, or at least a significant portion of a theory.

Holism is even more obvious if one considers sentences which are deeply embedded in the more theoretical realms of science, such as sentences about subatomic particles or about the most general features of space-time. Think, for example, of Einstein's famous statement that $E=MC^2$. The truth of that claim can only begin to be ascertained once one has some specification of what to count as energy, mass, and the speed of light, and how to identify and measure those things. Hence, a number of additional sentences, most of them also theoretical in nature, have to be brought in for the original one to imply a synthetic observation categorical and thus be tested. In fact, if we were to regard Einstein's statement in isolation, it could hardly be thought to have any meaning at all. Only in the context of a theory can we more or less specify what the sentence says and what has to be the case for it to be true. Strictly speaking, it is hard to see how the sentence itself, on its own, can be conceived as making a claim upon reality; only

the theory of which it is part, or more or less large portions of that theory, seem to make such a claim.

In ordinary scientific practice, a background theory is commonly presupposed each time terms such as 'energy' and 'mass' are used. Given those background presuppositions, theoretical sentences like Einstein's above may perhaps be thought to imply synthetic observation categoricals, even when considered in isolation from other sentences. Holism does not have to be thought as denying that. Rather, holism merely calls attention to those background presuppositions. It says that without them, theoretical sentences do not imply synthetic observation categoricals. Moreover, holism says that, in general, for those implications to obtain, the presuppositions built into the use of theoretical terms have to be fairly large in number and encompass a wide stretch of the theory or theories of which they are part. Assertion of a theoretical sentence oftentimes presupposes the assertion of a large portion of a theory, or at least of a system of interrelated sentences.[27]

One important and immediate consequence of holism is that if a theory implies an observation categorical that is disconfirmed by observations, no individual sentence of the theory is thereby immediately refuted. Rather, what is refuted is the theory, or system of sentences as a whole, together with whichever ordinary assumptions played a role in implying the failed observation categorical. Only theories as wholes, or portions of them large enough to imply observation sentences or synthetic observation categoricals, can be refuted by observations.

---

[27] Quine's holism is – in this regard at least – very much akin to a point later made by Wittgenstein: "What I hold fast to is not *one* proposition but a nest of propositions" (1969, § 225). For a comparative discussion of Wittgenstein and Quine on this particular issue, see Gibson (1996).

In the presence of an adverse observation, the scientist is in principle at liberty about which part of his theory to revise, and he may also choose to keep the theory as is and revise some ordinary assumptions which may have played a role in implying the false categorical. But he is of course constrained by the weight and strength that he assigns to various portions of his theory and ordinary assumptions, and by certain maxims of theory construction (such as simplicity, conservatism, generality, and fecundity). In general, there will be no choices at all between theories and hypotheses, or very few choices. Yet, in the face of a counterexample to his favored theory, a scientist always must in principle decide between various courses of action, even if only one of the alternatives seems reasonable to him at that moment.[28] He may choose to retain the hypothesis under scrutiny and revise some of the assumptions of his experiment, or perhaps some more fundamental tenets of his theory. More frequently, he will retain the experimental results and the basic assumptions of his theory, and revise only the specific hypothesis he set out to test. These considerations are confirmed by the practice of experimental scientists, who typically devise a number of alternative explanations for why an experiment may not have come out as expected.

Some statements are of course more easily revised than others; others are not as easily given up. The general laws and principles of natural science, for example, are typically upheld for decades before being replaced. At the extreme are the sentences of elementary logic and mathematics. Because they are common to all branches of science, they provide some unification to the various disciplines, and can only be revised at the cost of making considerable

---

[28] In many cases, it is in fact rather inappropriate to speak of a decision. Oftentimes the scientist will not even say that there are alternatives. This does not contradict holism; it just shows that in many cases the background theory held at the moment is thought to be beyond reasonable doubt, thus ruling out alternatives which would otherwise be thought to be relevant. For discussion on this topic, see Norton (1993, 1994, and 2003c), and Massimi (2004).

adjustments throughout. They are thus only very seldom changed; in most cases they are never changed. Nevertheless, changes do occur, and the so-called "scientific revolutions" are typical examples.

These rather trivial remarks make up the core of the thesis of holism. But they were also the source of, or the motivation for, some controversial claims which we find in Quine's TDE:

> The unit of empirical significance is the whole of science. (p. 42)
>
> [A]ny statement can be held true come what may, if we make drastic enough adjustments elsewhere in the system. (…) Conversely, by the same token, no statement is immune to revision. (p. 43)

Most of the criticisms of Quine on holism that one finds in the literature aim at those two passages. Quine himself, however, would later describe those claims as a "needlessly strong" (TDR, p. 268) assertion of the thesis. They are true, he would add, in a "legalistic sort of way" (*ibid*.), but they divert attention from what is more fundamental. Holism is an alternative to the claim that the empirical content of a theory can be sorted out distributively among the sentences and terms that comprise it. To contravene that claim, it suffices to say that some sentences of a theory, on their own, lack empirical content. There is no need to further claim or suppose that empirical content attaches only to the whole of science. Clusters of sentences "sufficient to imply an observable effect of an observable experimental condition" (TDR, p. 268) can also be thought as having an empirical content, even if they do not comprise the whole of science. It is true that science is considerably integrated, and that some components, such as elementary logic and mathematics, are common to all branches.

> But we can appreciate this degree of integration and still appreciate how unrealistic it would be to extend a Duhemian holism to the whole of science, taking all science as the unit that is responsible to observation. Science is neither discontinuous nor monolithic. It is variously jointed, and loose in the joints in varying degrees. In the face of a recalcitrant observation we are free to choose what statements to revise and what ones to hold fast, and these alternatives will

> disrupt various stretches of scientific theory in various ways, varying in severity. Little is gained by saying that the unit is in principle the whole of science, however defensible this claim may be in a legalistic way. (EESW, pp. 314-315)[29]

In his writings after TDE, Quine would likewise maintain that, legalistically speaking, it remains true that any individual statement can be revised or abandoned without affecting the net empirical implications of the resulting theory, so long as adjustments are made elsewhere in the theory or background assumptions. In the actual practice of working scientists, however, a large number of statements are effectively shielded from revisions because of their centrality to the theories accepted at that moment, or because they are directly confirmed by a very large number of observations. Scientists only subject a few statements to scrutiny at any given moment, and this is crucial for the development of the scientific enterprise. Revisions in statements that are central to the whole of science require a very large number of adjustments throughout, and alternatives to those statements are oftentimes unavailable at a given moment. These difficulties counsel against revisions, whenever possible. Quine offers these considerations as an explanation of the perceived 'necessity' of some theoretical statements; for example, those of mathematics:

> …mathematics infiltrates all branches of our system of the world, and its disruption would reverberate intolerably. If asked why he spares mathematics, the scientist will perhaps say that its laws are necessarily true; but I think we have here an explanation, rather, of mathematical necessity itself. It resides in our unstated policy of shielding mathematics by exercising our freedom to reject other beliefs instead. (*PT*, p. 15)

Revisions in those central statements occur only very seldom in the history of science, and while they occur other statements are maintained as they are.[30] Revisions of statements

---

[29] See also *PT*, p. 18, where Quine acknowledges the varying degrees to which a sentence may be subject to empirical testing: "In general (1) [Sodium chloride dissolves in water] is accepted as a vague statement of strong probability, open to question only where the improbable counter-instance can be plausibly accounted for. Similar cushioning shields much of science, it would seem, from the simple probe of observation categoricals."

directly linked to observations also require a large number of adjustments and are likewise shielded from revisions. In principle, or legalistically speaking, they are possible and have happened in the history of science.[31]

Quine replaced his "needlessly strong" assertion of holism in TDE with a more modest one in later works, which stresses what he calls the "empirical bias" of science, on the one hand, and the centrality of some theoretical sentences to the system of science as a whole:

> … the scientist does occasionally revoke even an observation statement, when it conflicts with a well attested body of theory and when he has tried in vain to reproduce the experiment. But the Duhem thesis would be wrong if understood as imposing an equal status on all the statements in a scientific theory and thus denying the strong presumption in favor of the observation statements. It is a bias that makes science empirical. (EESW, p. 314)[32]

Once these qualifications to the thesis of holism are granted, the thesis turns out to be quite modest, and its justification rather straightforward. Nonetheless, it has far-reaching consequences, not the least of which is the inseparability of questions of meaning from empirical questions, and the difficulties it raises for the traditional notion of 'a priori'.[33]

A stronger version of the thesis, criticized by Grünbaum (1960 and 1962) and suggested by the controversial passages in TDE quoted above, says that in the face of adverse observations

---

[30] A more or less clear illustration of this point can be found in the revision of our notions of space and time brought about by contemporary physics. The classical notions of space and time were indeed very central to modern science and philosophy; they were sometimes described as necessary and a priori notions – for example, by Kant – and revisions of them were thought to be absurd or impossible to philosophers up to the early nineteenth century. In 1827, Ferdinand Möbius, for example, rejected the idea of spaces with more than three dimensions because, he wrote, they "cannot be thought" (1991, p. 40-41). Spaces with more than three dimensions are nowadays routinely posited by string theories, for example.

[31] An example is belief that the Sun moves around the Earth. That belief was, in Ancient times, thought to be rather closely tied to the observation of the Sun moving across the sky every day.

[32] See also *FSS*, chapter 4.

[33] For further discussion, see Harman (1996 and 2003), and Hylton (2002 and 2004).

any statement can be held true by revising the accepted theory and ordinary assumptions so that the threatened hypothesis together with the revised theory will *imply* the observation categorical whose negation was implied by the original theory. Quine explicitly acknowledges Grünbaum's criticism, granting him the point while maintaining his own version of the thesis: "Inactivating the false implication is all that is at stake." (*PT*, p. 16)[34]

Quine's holism merely says that given an hypothesis *H*, an accepted theory *T*, and a synthetic observation categorical *O* which is found to be false, if (*H* and *T*) implies *O*, then one can revise *H* and *T* and come up with *H′* and *T′* such that (*H′* and *T′*) does not imply *O*. Moreover, one should also be able to come up with at least one alternative revision, say, *H″* and *T″*, such that (*H″* and *T″*) does not imply *O* either. The stronger version of holism criticized by Grünbaum says instead that in the face of adverse observations *H* and *T* can be revised so that (*H′* and *T′*) implies not-*O*, and likewise for (*H″* and *T″*). The latter, stronger version of holism is not trivially true, since alternative theories that imply a given observation are not always easy to think up and it is not clear that they must exist for all cases. Oftentimes, in the history of science, some observations remain anomalous for considerable periods of time. Coming up with theories that can predict and explain such anomalies is not asserted by Quine's thesis, which merely says that alternative ways of *preventing* a theory from entailing a false observation categorical are always available. To prevent an implication from obtaining it suffices to weaken one or more sentences of the accepted theory or of the hypothesis under consideration. Whereas "[e]xplaining the unexpected counter-observation is quite another step of scientific progress, which may or may not be made in the fullness of time" (*PT*, p. 16).

---

[34] See also CGC, p. 132, and TI, pp. 11-12. Laudan (1990, p. 271 ff.), Kitcher (1993, p. 250), and Leplin (1997, p. 210), however, have insisted in assigning to Quine the stronger version of holism criticized by Grünbaum, despite Quine's explicit disavowals.

2.3 **Duhem and Carnap**

Prior to Quine, the thesis of holism had already been stated by Duhem:

> A physicist decides to demonstrate the inaccuracy of a proposition (…) The prediction of the phenomenon, whose nonproduction is to cut off debate, does not derive from the propositions challenged if taken by itself, but from the proposition at issue joined to [a] whole group of theories; if the predicted phenomenon is not produced, not only is the proposition questioned at fault, but so is the whole theoretical scaffolding used by the physicist. The only thing the experiment teaches us is that among the propositions used to predict the phenomenon and to establish whether it would be produced, there is at least one error; but where this error lies is just what it does not tell us. (Duhem, 1954 [1914], p. 185)

Duhem's thesis has a narrower scope than Quine's, however. Duhem takes holism to be true of theoretical hypotheses in physics and a few other areas of natural science that are sufficiently removed from ordinary experience; but he takes it to be false of mathematics and logic, on the one hand, and of physiology, certain branches of chemistry, and other disciplines which are more closely tied to observations than theoretical physics. In these disciplines, he wrote, it is expected of the scientist that he "establish an absolute separation or watertight compartment between the consequences of his theoretical deductions and the establishing of the facts shown by his experiments" (1954, p. 182), whereas in physics such separation cannot exist: "it is impossible to leave outside the laboratory door the theory that we wish to test, for without theory it is impossible to regulate a single instrument or to interpret a single reading" (*ibid*.). Moreover, in physics the instruments used in experiments only perform their specific functions because a certain *physical* explanation of the way they work is presupposed. Thus, in theoretical physics, Duhem concludes, only "whole theoretical groups" can be tested, isolated hypotheses cannot.

Duhem argues that, contrary to what happens in theoretical physics, the results of experiments in physiology, for example, do not presuppose a theoretical (physiological)

explanation of the workings of the equipment used in the experiments. Experiments in physiology can thus be regarded as establishing facts regardless of what theories in physiology might lead one to expect as the right outcome of those experiments. Observations would in that sense override theoretical hypotheses. Similarly, but on the other end of the spectrum, in mathematics and logic, a strict separation also exists between theory and observation. In this case, however, observations would be completely immaterial: in establishing the truths of mathematics and logic, experiments do not matter.

While reading Duhem nowadays, one is forcefully reminded of changes that chemistry and physiology, as well as math and logic, have undergone in the century that has passed since he wrote his work. Nowadays, chemistry and physiology are as theoretical as almost all other branches of natural science, and the separation Duhem speaks about has become much harder to detect, if it exists at all. Also, physics seems now even more tightly connected to the developments in mathematics and logic than it was in Duhem's time. As Quine would often note (for example: TDE, p. 43), alternative logics have been proposed as ways of dealing with empirical problems brought about by quantum mechanics.

Quine's holism is, accordingly, broader in scope than Duhem's: it purports to hold of all branches of science, including mathematics and logic. It is likely that this difference between Duhem and Quine is at least in part due to the scientific developments that occurred in the first half of the twentieth century, which only Quine witnessed.[35] Duhem must have had in mind the science of his time, which was a lot less theoretical than it later became. Quine himself also connects the difference between his version of the thesis and Duhem's to his own "view of

[35] But it is also likely that Quine's background as a logician has a strong influence on his views on these matters: he tends to address them from a very abstract and general level.

common sense as primitive scientific theory" (RJV, p. 619) and his realistic attitude towards the theoretical entities and principles posited by science, as opposed to Duhem's fictionalistic attitude. Duhem, like some others in the first half of the twentieth century, thought that the physical hypotheses positing theoretical entities and principles were strictly speaking neither true nor false, but symbolic approximations intended to facilitate or produce adequate predictions of observations.[36] Quine, on the other hand, thought of the posits of science as on a par with directly observable objects posited by common sense, such as chairs and tables: "Science is a continuation of common sense, and it continues the common-sense expedient of swelling ontology to simplify theory." (TDE, p. 45).[37]

Duhem originally developed his version of the thesis of holism in the late nineteenth and early twentieth century. Quine's views were developed independently in the 1930s and 1940s,[38] as a reaction to his reading of Carnap's *Aufbau*.[39] On Quine's reading, Carnap fully embraced a reductionist project in the *Aufbau*, which aimed at translating, and thus reducing, all of science into logic and observation terms (EN, p. 76). Quine praises Carnap as "the first empiricist who,

---

[36] The following passage by Einstein illustrates this point: "[The general theory of relativity] revealed that is was possible for us, using basic principles very far removed from those of Newton, to do justice to the entire range of the data of experience in a manner even more complete and satisfactory than was possible with Newton's principles. But quite apart from the question of comparative merits, the *fictitious character* of the principles is made quite obvious by the fact that it is possible to exhibit two essentially different bases, each of which in its consequences leads to a large measure of agreement with experience. This indicates that any attempt logically to derive the basic concepts and laws of mechanics from the ultimate data of experience is doomed to failure." (1934, p. 166 – emphasis added)

[37] For a comparative analysis of Duhem's and Quine's theses, see Vuillemin (1998) and Quine's reply (RJV).

[38] Commenting back on TDE, Quine wrote: "In a footnote to 'Two Dogmas' I noted Duhem's priority in stressing holism. As a matter of curiosity, however, I might mention that when I wrote and presented 'Two Dogmas' … I didn't know about Duhem. Both Hempel and Philipp Frank subsequently brought Duhem to my attention, so I inserted the footnote when 'Two Dogmas' was reprinted in *From a Logical Point of View*." (TDR, p. 269)

[39] See Carnap (1967).

not content with asserting the reducibility of science to terms of immediate experience, took serious steps toward carrying out the reduction" (TDE, p. 39). [40]

The *Aufbau*, however, contains only a very rough guide on how to proceed in order actually to produce a reduction of science into logic and observation terms. Yet, even apart from sketchiness, the project as laid down by Carnap turned out to be unviable. Only a few years after its publication Carnap had already substantially weakened it.[41] On Quine's reading, the project failed because it could not fulfill its promise of "specifying a sense-datum language and showing how to translate the rest of significant discourse, statement by statement, into it" (TDE, p. 39).[42] The crucial problem is made evident in sections 126 and 127 of the *Aufbau*, where Carnap lists a set of desiderata that need to be fulfilled in order for the reduction to take place. Those desiderata concern the reduction of sentences expressing perceptions of colors to sentences assigning colors objectively to space-time points. The sections are meant to illustrate how the reduction of scientific sentences in general is supposed to take place, and they mark what Carnap calls the passage from the "autopsychological" realm to the objective world. They attempt to show how solipsistic, or subjective, constructions of sensations and experiences can be assigned objective space-time points. By specifying a recipe, however sketchy, for making those assignments, it would show how to reduce all statements about color properties to statements comprising only logical terms and terms for sense-data, and provide a guide for the assignment of other types of sensations and experiences to the objective world. Carnap lists 11 desiderata in

---

[40] Quine's reading of the *Aufbau* has been contested in the recent literature: see Michael Friedman (1999) and Alan Richardson (1998). These authors emphasize Carnap's neo-Kantianism: the *Aufbau*, on their reading, was an attempt to explain all knowledge in logical terms only, without the pure concepts of the understanding which were at the heart of Kant's metaphysics.

[41] See Carnap (1936 and 1937), and EN, p. 77; see also Carnap's Preface to the second edition of the *Aufbau* (1967, p. viii).

[42] See also EN, pp. 76-77.

§ 126, which are subsequently supplemented (§ 127 contains an alternative presentation of the same material).

The idea is to assign each color sensation to a certain space-time point, and assign colors to all the remaining space-time points so as to maximize the overall continuity and simplicity of the system of space-time points, at least inasmuch as colors are concerned. If, for example, the color red is assigned to point *p* at time *t*, and the same assignment is made a few minutes later, then in the time interval in between those two assignments we should assume that point *p* remained red, unless some other assignments determine otherwise. Thus, desideratum 10 says that

> [W]e have to assign a color to the unseen color spots. Taking into account the colors of seen color spots, we make a preliminary choice of these colors in such a way that the color of the points of a world line, considered as a function of time, shows a rate of change which is as small as possible, i.e., if possible, remains constant. (1967, pp. 196-197)

Likewise, as new sensations are had and new assignments are made, previous assignments may need to be revised accordingly, always attempting to maximize the overall continuity and simplicity of the assignments. Since the desiderata listed by Carnap are meant to provide a general recipe for specifying which color sensation *is at* each space-time point, they can be understood as providing instructions for reducing the relation "is at" to logical terms and terms for sense data. As Quine points out, however, the intended reduction cannot obtain even in principle:

> Carnap did not seem to recognize (…) that his treatment of physical objects fell short of reduction not merely through sketchiness, but in principle. Statements of the form 'Quality *q* is at point-instant *x*;*y*;*z*;*t*' were, according to his canons, to be apportioned truth values in such a way as to maximize and minimize certain over-all features, and with growth of experience the truth values were to be progressively revised in the same spirit. I think this is a good schematization (deliberately oversimplified, to be sure) of what science really does; but it provides no indication, not even the sketchiest, of how a statement of the form

> 'Quality *q* is at *x*;*y*;*z*;*t*' could ever be translated into Carnap's initial language of sense data and logic. The connective 'is at' remains an added undefined connective; the canons counsel us in its use but not in its elimination. (TDE, p. 40)

Furthermore, Carnap does not give any indication of the relative weight of each desiderata. Hence, conflicting desiderata may in principle yield conflicting assignments of colors to objective of space-time points. More importantly for our purposes here, the assignment of colors to space-time points cannot be executed one at a time, since some assignments are dependent on others. If a certain sensation of color is an hallucination, for example, there is no corresponding assignment to be made in the objective space-time. But whether a sensation is an hallucination can only be judged by comparing it to other sensations that are not hallucinatory. Hence, the assignments of colors to objective space-time points cannot be made one by one. Rather, they must proceed from the totality of sensations (or a very large number of them) to the totality of objective space-time points (or a large portion of them). However, since the totality of color-sensations is never completely given, one can never be sure that the assignments made up to a given time will not have to be revised later in the light of new sensations. Thus, one can never actually reduce, one by one, the sentences containing assignments of color to objective space-time points to a sentence containing only terms for sense-data and logical terms.

To his credit, this is a point which Carnap seems soon to have acknowledged, since in later writings he proposes a reduction of a laxer kind. In "Testability and meaning" (1936 and 1937), for example, he puts forth the weaker notion of 'reduction form' as an alternative to the explicit definitions proposed as instruments of reduction in the *Aufbau*. Perhaps not surprisingly, in the *Logical Syntax of Language* [1937], Carnap explicitly endorses Duhem's holism:

> … it is not possible to lay down any set rules as to how new primitive laws are to be established on the basis of actually stated protocol-sentences. (…) Further, it is in general impossible to test even a single hypothetical sentence. In the case of a

> single sentence of this kind, there are in general no suitable L-consequences of the form of protocol-sentences; hence for the deduction of sentences having the form of protocol-sentences the remaining hypotheses must also be used. Thus *the test applies, at bottom, not to a single hypothesis but to the whole system of physics as a system of hypotheses* (Duhem, Poincaré).
>
> No rule of the physical language is definitive; all rules are laid down with the reservation that they may be altered as soon as it seems expedient to do so. This applies not only to the P-rules but also to the L-rules, including those of mathematics. In this respect, there are only differences in degree; certain rules are more difficult to renounce than others. (2002, pp. 316-317)

Incidentally, Quine lectured on the philosophy of Carnap in 1934 (see his LC), and may have picked up the gist of the reasons for holism from Carnap himself.

2.4 **Holism lends credence to underdetermination**

The thesis of holism, even in its more modest version defended by Quine after TDE, suggests the plausibility of a different one, the thesis of underdetermination. Quine changed his formulation of the thesis of underdetermination a couple of times (see Chapter 3). We shall be concerned here with his final formulation of the thesis, which says that there are empirically equivalent rival systems of the world which, if we were to find them, we would find no way of rendering logically equivalent through reconstrual of predicates, or translation. This latter clause intends to rule out theories that are rivals only in a linguistic or terminological sense. For underdetermination to hold, there must be (in a Platonic sense) theories that are rivals and which we would be unable to render logically equivalent through any known manual of intertranslation. Thus formulated, the thesis of underdetermination is suggested by holism:

> If in the face of adverse observations we are free always to choose among various adequate modifications of our theory, then presumably all possible observations are insufficient to determine theory uniquely. (EESW, p. 313)

Hence, the suggestion goes, presumably there are systems of the world which are empirically equivalent and yet rivals in the relevant sense. The reasoning presented in the

passage just quoted seems clear enough at first glance: each "adequate" modification would yield a new theory which is compatible with observations in that it does not imply observation categoricals that are known to be false. Since various new theories are thus made possible, the observations available at any given moment cannot, on their own, determine one theory above all others. This, however, is not yet the thesis of underdetermination, but just an immediate consequence of holism itself.[43] As Quine points out in the passages that follow the one quoted above, the thesis of underdetermination further enjoins us to entertain the possibility of rival theories that imply exactly the same observation categoricals and yet differ from one another at the level of theoretical sentences. As such, the thesis of underdetermination cannot be thought to follow from holism alone, for two main reasons:

(1) Given holism, we have reason to expect that various theories can be designed to conform to a given set of observations. "Adequate modifications" can yield new theories which are consistent with observations in the sense that they do not entail observation categoricals that are known to be false. However, those "adequate modifications" of a theory could, for all we know, each imply a different, or perhaps a new set of observation categoricals. The revised theories, in other words, might not end up being empirically equivalent. In fact, that is to be expected, since revisions of different hypotheses are likely to affect the implied observation categoricals in different manners. Hence, even though the new theories may each be compatible with the observations that led to the revisions of the original theory, they may not be compatible with future observations, or past events that went unobserved. Hence, they fail as cases of underdetermination.

---

[43] Some authors – for example, Sklar (1975 and 1981) and Stanford (2001) – refer to this thesis as "transient underdetermination". It is not clear, however, how that thesis differs from holism. See Hoefer and Rosenberg (1994) for discussion of this particular point.

(2) Even if the new theories, produced by "adequate modifications" of an older theory, were all empirically equivalent, one cannot thereby conclude that they are in fact distinct theories in any interesting sense. This of course depends on which criteria for theoryhood one adopts. But even in the absence of clear criteria of theoryhood, there a question one can raise independently: Since empirically equivalent theories must by definition agree on all observation categoricals, whatever differences they may have must be confined to sentences that hinge on observations only indirectly, that is, the so-called "theoretical sentences". Thus, one cannot in principle exclude the possibility that all the new theories produced by the revision process are not only empirically equivalent but in fact theories which can be rendered logically equivalent by "reconstrual of predicates", or translation.[44] A physical theory in which all occurrences of "proton" and "electron" are interchanged is indeed empirically equivalent to the original theory in which those terms are not interchanged. The two theories, however, can be clearly rendered logically equivalent by translating both into the vernacular English of technical physics. This is a fairly trivial case. In principle, however, more complicated cases might also turn out to be likewise intertranslatable, even if a manual of translation is hard to come by.

Holism is thus insufficient to establish underdetermination.[45] The latter can only obtain if not only alternative ways of revising a theory can render it compatible with a given set of observations, but also if the resulting theories are empirically equivalent but cannot be rendered logically equivalent through intertranslation. There are considerable difficulties in demonstrating that this is possible even in principle, and for that reason Quine tended to treat the thesis as a conjecture (see, e.g., *TT*, p. 181). Whether there are empirically equivalent systems of

---

[44] The issue of intertranslatability of theoretical sentences is discussed in more detail in chapter 4.

[45] Quine is frequently misread on this point: Laudan (1990, pp. 271 ff.) and Kitcher (1993, p. 251), for example, suggest that for Quine underdetermination is "grounded" on holism.

the world that are not intertranslatable, he wrote, is an "open question" (EESW, p. 327). Also open is the question whether, upon finding such theories and failing to intertranslate them, that failure is to be accounted on our lack of ingenuity or on intrinsic features of the two theories that would prevent any intertranslation in principle (EESW, p. 328). Underdetermination, in other words, although plausible on a first approximation, cannot be asserted as categorically as holism can.

2.5 **Some implications of the two theses**

Within Quine's philosophy, holism has far-reaching consequences, despite its trivial justification. Quine argued that it provides good reasons to reject the idea of a realm of epistemologically privileged sentences – say, "analytic" or "a priori" sentences – which would be immune to revision or true in virtue of meaning or necessarily true. All the sentences that make up the corpus of a scientific theory are justified by the net empirical implications of that theory. If the synthetic observation categoricals implied by the theory are confirmed, so are all the sentences that comprise the theory, however removed they might be from observations. The thesis thus entails that the traditional distinctions between the various branches of science must be quite plastic, and that even the traditional distinctions between metaphysics and science, and between empirical claims and theoretical claims are rather arbitrary. Holism thus supports some substantial theses which are at the core of Quine's philosophy.

Quine also used the thesis of holism in his justification the thesis of indeterminacy of translation:

> [I]f the English sentences of a theory have their meaning only together as a body, then we can justify their translation into Arunta only together as a body. There will be no justification for pairing off the component English sentences with

> component Arunta sentences, except as these correlations make the translation of the theory as a whole come out right. Any translations of the English sentences into Arunta sentences will be as correct as any other, so long as the net empirical implications of the theory as a whole are preserved in translation. (EN, p. 80)

Indeterminacy of translation, like underdetermination, is rather controversial thesis in the secondary literature. As is the case with underdetermination, it has also been frequently misunderstood. Besides unnerving commentators, the two theses share another interesting parallel: neither categorically affirms its peculiar form of indeterminacy. Quine distinguishes the thesis of indeterminacy of translation, from ontological relativity (or indeterminacy of terms). The latter, he wrote, admits of trivial proof: "The essence of the proof is just that *x* is an *F* if and only if the proxy of *x* is the proxy of an *F*" (RJW, p. 728). The former, on the other hand, is a conjecture.[46] In this regard the thesis of underdetermination is similar to the thesis of indeterminacy of translation (or indeterminacy of sentences, or "holophrastic meaning").

The thesis of underdetermination, because of its conjectural status, cannot, on its own, support the weight of substantial doctrines. This is a point on which Quine is widely misread. The following passage, by Laudan and Leplin (1991), illustrates well the misunderstanding in question:

> By the 1920s, it was widely supposed that a perfectly general proof was available for the thesis that there are always empirically equivalent rivals to any successful theory. (…) [B]y the 1940s and 1950s, it was thought that – in large part because of empirical equivalence – theory choice was radically underdetermined by any conceivable evidence. Whole theories of knowledge (e.g., W. V. Quine's) have been constructed on the presumption that these results are sound. (p. 449)

It is certainly true that holism (alongside empiricism and naturalism) plays a fundamental role in Quine's theory of knowledge. Underdetermination, given its rather conjectural justification, cannot play such role. The thesis suggests that alternative descriptions of the world

---

[46] See RJW (p. 728), OR (p. 33), and TI (pp. 4-7).

may turn out to be equally tenable. But this is not an implication which can be asserted categorically within Quine's philosophy. Rather, the possibility of rival theories of the world which are empirically equivalent and yet non-intertranslatable remains "an open question" (EESW, p. 327). Given holism and the "less-than-rigid" connections that seem to obtain between theories and observations, it is to remain a plausible conjecture, however.

In EESW, Quine takes underdetermination to hold an important lesson for our attitudes towards science: however certain we may be of the correctness of the theories we hold, however well those theories may be supported by observations, there may always be room for "undiscovered systematic alternatives" (EESW, p. 327) which may remain forever undetected. Symptomatically, the lesson Quine derives from underdetermination concerns our *attitudes* toward science, and not any metaphysical or epistemological doctrine. As we shall see in the following chapters, this is in keeping with Quine's deflation of the metaphysical problems associated with the thesis; Quine's thesis of underdetermination is systematically at odds with stronger versions of the thesis currently held in the literature.

3. QUINE'S VIEWS ON UNDERDETERMINATION

As mentioned above, Quine is often regarded as one of the main proponents of the thesis that natural science is underdetermined by observations.[47] His views on the matter, however, changed several times over the years. He has maintained one version or another of the thesis throughout his writings, but his formulations were revised at least twice, and at least three times he changed his mind about whether empirically equivalent yet rival systems of the world can be simultaneously thought to be true.[48] His views on underdetermination are thus not easy to apprehend, at least on a first reading, and perhaps for that reason have been frequently misrepresented in the literature.[49]

A common mistake consists in assigning to Quine belief in a thesis stronger than the one he actually put forth. This is not surprising. Given his influence as a proponent of the thesis, one is rather naturally led to believe that it must have been a significantly strong thesis. Hence, it is also natural to believe that the thesis must have been a somehow fundamental in his philosophy. Neither is the case, however. The version of the thesis which Quine eventually settled for is too weak to support any strong metaphysical or epistemological doctrines, and it is neither obviously true nor is it immediately clear what its implications are. It is nonetheless a plausible thesis,

[47] See, for example, Sklar (1975, p. 379), Newton-Smith (1978, p. 71), Horwich (1982, p. 61), Worrall (1982, p. 202), Ariew (1984, p. 313), Bergström (1984, p. 349; 1993, p. 331; 2004, pp. 91ff.), Ben-Menahem (1990, p. 262), Laudan (1990, pp. 271 ff.), Laudan and Leplin (1991, p. 449), Earman (1993, p. 31), Kitcher (1993, pp. 249-251), Hoefer and Rosenberg (1994, pp. 593 f.), Kukla (1996, p. 139), Leplin (1997, p. 203), Yalçin (2001, *passim*), Stanford (2001, p. S8 n.), Devitt (2002, pp. 31 ff.), Okasha (2002, p. 304), and Massimi (2004, p. 243). In fact, nearly all recent publications on underdetermination acknowledge Quine's contributions.

[48] See PR (1955, p. 254); *WO* (1960, pp. 22-23); RC (1969, p. 302); RIT (1970, pp. 178-179); EESW (1975, *passim*); NNK (1975, p. 79); TTPT (1981, pp. 21-22); EC (1981, pp. 29-30); RA (1984, p. 294); RRG (1986, pp. 156-157); ITA (1987, pp. 9-10); TI (1990, pp. 13-15); *PT* (1992, pp. 95-101); SN (1992, p. 9).

[49] A noteworthy exception is Gibson (1988 and 1998). See also Bergström (1990, 1993, and 2004).

especially for an empiricist such as Quine; moreover, it is suggested by other doctrines within his philosophy, holism in particular, as we saw in chapter 2.

The main goals of this chapter are to lay out Quine's views on underdetermination as clearly as possible, to track down the changes they underwent, and to try to understand why they have been so systematically misrepresented. This has been done only partially in the literature, most notably by Gibson (1988 and 1998).[50] The intention here is both to complement existing expositions and correct some common misreadings. Some implications of Quine's analysis of the thesis for the ongoing debates on the matter are explicitly stated. A secondary goal is to indicate that despite Quine's influence as a proponent of the thesis, his views contain a rather neglected alternative to the ones which currently prevail in the literature.

The chapter is divided as follows: section 3.1 briefly introduces the issues; section 3.2 tracks down changes in Quine's formulation and justification of the thesis up to 1975; section 3.3 discusses later changes and the corresponding vacillations on the truth of empirically equivalent rival theories; and section 3.4 concludes with a summary and brief discussion of some of the ways in which Quine's views have been misunderstood.

3.1 **Introduction**

As mentioned above, Quine consistently upheld the thesis of underdetermination but changed his mind several times about its consequences and about how to best formulate it. While asserting the thesis, he often qualified it with phrases which indicate that he did not regard it as immediately clear or evident. He wrote, for example, that underdetermination is "slippery

[50] See also Bergström (1990, 1993, and 2004).

when we try to grasp it more firmly" (NNK, p. 80) and "plausible insofar as it is intelligible, but less readily intelligible than it may seem" (EESW, p. 313). None of this seems too characteristic of Quine's work or style. This was a philosopher who throughout his life changed position very little about his core doctrines (holism, naturalism, empiricism), and tended to assert them quite forcefully. It is also unexpected that as someone identified as a proponent the thesis, he should be somewhat hesitant to assert its intelligibility and vacillate about some of the consequences of the thesis. Thus, it might not be surprising that his qualms on these matters have been repeatedly interpreted as a symptom of a deeper tension within his philosophy, a tension between his naturalism and his unrelenting empiricism.[51]

Also unexpected for someone regarded as one of the main proponents of the thesis is the fact that Quine's views differ significantly from almost all others that have been recently published. Debates on underdetermination have been taking place quite independently of Quine's writings, despite his acknowledged influence on its original formulations. In the recent literature, two major attitudes towards underdetermination stand out: (i) Laudan and Leplin (1991), Kitcher (1993), Norton (1994 and 2003c), Massimi (2004), and others, challenge both the idea that rival theories can be empirically equivalent and the perceived entailment of underdetermination from empirical equivalence (see Chapter 5 for further explanation and discussion). According to those authors, currently held arguments for underdetermination rely on an oversimplified account of scientific practice, which is to be overcome by closer attention to the role of auxiliary assumptions and to actual historical cases of alleged empirical equivalence. (ii) On the other hand, Van Fraassen (1980), Bergström (1993), Kukla (1996), Stanford (2001),

---

[51] Gibson (1986), Bergström (1993), and Yalçin (2001), for example, maintain that it is a symptom not merely of a tension in his philosophy but of an outright inconsistency – although Gibson is less critical in a later work (1988, pp. 113-124).

and others, seem confident that underdetermination can be established (either by developing empirically equivalent rivals to any given theory by an appropriate algorithm, or by inductive reasoning over the history of science). They maintain, accordingly, that because of underdetermination, we should back down on at least some aspects of scientific realism: either theories (or at least their non-observable consequences) ought not to be thought as attempting to identify and describe reality but rather as tools for reliable predictions of observations, or they are to be thought as attempting to identify and describe reality but we are to remain always agnostic about whether that goal has been achieved. Attitude (i) is the one prevailing among epistemic *realists*. Attitude (ii) prevails among epistemic *anti-realists* and agnostics.[52]

However, all the above mentioned authors have in common fact that they perceive underdetermination as somehow incompatible with scientific realism. Underdetermination says that alternative, rival theories of the world can be equally warranted by observations. Quine describes those theories as "empirically equivalent", and what he means is that they imply the same observation categoricals (see *PT*, pp. 16-17 and 95-96). Since they are rivals, they are presumably irreducible to one another. Hence, the set of entities and principles posited by each must differ. Thus the world according to one theory would not be the same as the world according to its empirically equivalent rivals. Scientific realism enjoins us to interpret scientific theories literally, and to assume that the entities and principles posited by such theories are real, insofar as those theories are thoroughly confirmed by observations. Hence the conflict:

---

[52] Yet another view found in the recent literature is that of Sklar (2000), who focuses his attention not on philosophical theories or abstract methodological accounts of science but on actual scientific practices instead. His claim is that some of the concerns that motivate underdetermination can already be found in actual scientific reasoning, and not just in philosophical reflection about that reasoning. He also claims that scientists typically refrain from drawing some of the broader philosophical consequences that one often finds in more abstract reflections about the methodology of science. In this regard, as we shall see, Sklar's views have important similarities with Quine's.

empirically equivalent theories seem to stand on the same footing in regards to warrantedness; if we are to assign truth to one, there seems to be no reason to refrain from assigning truth to the others as well. But if those theories are taken at face value, as scientific realism enjoins us, it seems that at most one of any two conflicting theories can be true. It is thus thought that if underdetermination holds, either we must suspend judgment as to the truth of empirically equivalent but conflicting theories, or maintain that they do not literally depict reality but are merely tools for predicting sensory stimuli. Reactions to the thesis, accordingly, tend to fall into the two groups mentioned above: either (i) the thesis of underdetermination is false or confused, or (ii) it entails anti-realism or agnosticism.[53]

Quine's views on underdetermination differ from both (i) and (ii). He tends to downplay the theoretical significance of underdetermination: "the more closely we examine [the thesis] (…) the less we seem to be able to claim for it a theoretical basis" (EESW, p. 326), and seems quite reluctant to assign to it any strong metaphysical or epistemological consequences. In fact, Quine did maintain the thesis while holding fast to realism, which by itself already indicates a divergence with the two attitudes described above. As we shall see below, there was quite a lot of vacillation in Quine's own views. But he progressively mitigated the thesis, and in one of his last published treatment of the matter, he came to suggest that some of the considerations which motivate attitudes (i) and (ii) above would "simmer down, bathetically, to a question of words": "The fantasy of irresolubly rival systems of the world is a thought experiment out beyond where linguistic usage has been crystallized by use." (*PT*, pp. 100-101)

---

[53] For further discussion, see chapter 5.

3.2 **"The doctrine is plausible insofar as it is intelligible"**

Quine's most detailed and thorough analysis of underdetermination is his 1975 EESW. Prior to that, he discussed the thesis only very briefly in various passages, such as PR (p. 254), and *WO* (p. 23). In those earlier texts, he neither argues for the thesis extensively nor goes into much detail about what it amounts to or how it can be spelled out. Rather, he seems to rely merely on its overall plausibility, which he takes to follow from the way the statements of natural science relate to observations: In a fully developed and theoretically complex scientific theory, such as the ones currently held in physics, there is always a gap between scattered even if methodically collected observation reports, on the one hand, and the theoretical principles and entities which systematize those reports into a theory, on the other. In a typical modern physical theory, correct predictions of a large number of observable events can only be obtained when theoretical principles and entities are posited alongside observable entities. Evidence for those theoretical principles and entities can only be indirect, since it comes from the role they play in observable predictions. The connections that thus obtain between observations and the more theoretical portions of a theory remain always "less-than-rigid" (PR, p. 254), which makes it plausible to think that alternative sets of theoretical principles and entities can be devised that allow for exactly the same predictions of observations.[54]

Quine's reasoning in PR and *WO* roughly follows the lines just sketched. The plausibility of underdetermination in Quine's thought is further strengthened by his description of scientific practice as roughly conforming to the so-called 'hypothetico-deductive method', at

[54] See *WO*, § 6; TTPT, *passim*; NNK, pp. 74-75, 79; *PT*, chapter 1.

least in regards to the testing of scientific statements.[55] Clusters of statements imply observation categoricals, which are then compared to observations. Alternative clusters of statements may in principle imply the same observation categoricals, and indeed do so with equal theoretical simplicity or economy.[56] Contrary to much of what has been said of Quine in the recent literature, his reasoning on these matters is only conjectural, and his conclusion is not a categorical assertion of a positive thesis, but almost a concession in the absence of good reasons to think otherwise:

> … we have no reason to suppose that man's surface irritations even unto eternity admit of any one systematization that is scientifically better or simpler than all possible others. It seems likelier, if only on account of symmetries or dualities, that countless alternative theories would be tied for first place. Scientific method is the way to truth, but it affords even in principle no unique definition of truth. (*WO*, p. 23)

Rival, alternative theories could then be legitimately described as "empirically equivalent": Any evidence for one such theory would likewise corroborate the others.[57] They

---

[55] Norton (1994), Massimi (2004) and others have argued that the hypothetico-deductive method provides at best an impoverished representation of actual scientific practice, which leaves unexplained the certainty with which currently accepted theories are upheld against relevant rival alternatives by scientists. They also argue that part of the plausibility of the thesis of underdetermination comes precisely from such an impoverished representation. In part, the point is well made against Quine. But we should note that his main concern was with formulating a *general* description of the relations between theories and observations. At that level of abstraction, Quine can insists that the hypothetico-deductive method captures the essence of scientific practice, even if the construction of new theories and the choice between rival theories is, in practice, constrained, as Norton and Massimi point out, by currently held hypotheses and practices in neighboring areas of science. Given such constraints, a choice may even seem to be determined by the available evidence. This is explicitly acknowledged by Quine in various passages, for example, *PT*, p. 18.

[56] This kind of reasoning is of course not original in Quine. As we saw in Chapter 2, Pierre Duhem was already arguing in a similar vein several decades earlier: "Shall we ever dare to assert that no other hypothesis is imaginable? Light may be a swarm of projectiles, or it may be a vibratory motion whose waves are propagated in a medium. Is it forbidden to be anything else at all?" (1914, p. 190) See also Poincaré (1905, chapter 10, especially pp. 167 ff.).

[57] For Quine, two theories are empirically equivalent if they entail the same observation categoricals (EESW, p. 319; EC, p. 28; TI, p. 13). He seems to adopt here a view that has been very much criticized in the recent literature, suggesting that all evidence in favor of or against a theory has to fall within the scope of the observation categoricals that are entailed by the theory. Against that view, Laudan and

would agree on all observation categoricals[58] and thus make exactly the same observable predictions. Yet they would diverge on the unobservable entities each posits or on the theoretical principles each asserts.

Room for divergences between empirically equivalent theories increases as science becomes theoretically more complex and sophisticated. Science has become an increasingly theoretical affair: currently held views on electrons and quarks, the principle of general relativity, and models for quantum theory, can provide some examples. For Quine, underdetermination can *only* obtain in such cases, where the internal structure of a theory is unavoidably complex and does not directly hinge on observations. Theories which posit no theoretical principles and

---

Leplin (1991), Massimi (2004), and others, have argued that evidence for a theory can also come from confirmed observation categoricals that are not directly entailed by the theory in question but by a more general theory, which in turn entails the theory in question. In such cases, direct evidence for the more general theory is also indirect evidence for the local theories that it entails. The point, although well taken, is only relevant where local theories are concerned. Quine's discussions of underdetermination, however, focus primarily if not exclusively on global theories, where indirect evidence coming from more general theories simply cannot obtain. His discussions are therefore conjectural, since we have never had an actual global theory of the world. The best candidate so far is physics, but even physics is yet to be unified. Quantum mechanics and general relativity are yet to be reconciled. Nevertheless, it is such a unified global theory that Quine has primarily in mind when discussing underdetermination. This is reasonably clear in several passages; for example: "Let us limit [the case] further to global systems of the world, so that there is no question of fitting the rival theories into a broader context." (*PT*, p. 98) Also, his use of the phrase "systems of the world" (EESW, p. 313; CNS, p. 66; RRG, p. 155; TI, p. 14; and *PT*, pp. 95-102) throughout his discussions indicates that he has global theories in mind.

[58] In texts written in the 1970s and before, Quine uses the notion of 'observation conditional' rather than 'observation categorical'. An observation conditional is a conditional sentence in which both the antecedent and the consequent are conjunctions of pegged observation sentences (EESW, p. 318). A pegged observation sentence is "a non-observation sentence obtained by pegging an observation sentence" (p. 328, n. 2) with space-time coordinates. Observation conditionals are thus standing sentences, "meant to be true or false independently of the occasion of utterance" (p. 316). In this respect, they differ from observation sentences, which command "assent on some occasions and not others, depending on what is happening where and when the sentence is queried" (*ibid.*). From the early 1980s on, Quine stopped using the notion of 'observational conditional', favoring instead that of 'observation categorical', which does not presuppose a space-time coordinate system, and is thus less theory-laden: "We can withdraw to what I may call *observation categoricals* – sentences like 'Where there is smoke there is fire' or 'When it rains it pours' or 'When night falls the lamps are lit'. These enjoy generality over places and times, but they do not need to be read as assuming a prior ontology of places and times or any implicit universal quantification over them." (EC, p. 27) The change in the notion favored by Quine is not significant for our present purposes here.

entities cannot be underdetermined, unless by "rival theory" one has trivial linguistic variations in mind. Theories indeed do have linguistic alternatives, but surely this is not what proponents of the thesis of underdetermination intend assert. If it were, then any theory written in in German and its English translation would count as alternatives, even if the translations were straightforward and unambiguous.

In EESW (p. 323) Quine argues that theories which entail only a finite number of observation conditionals should not be regarded as underdetermined either. He maintains that rival alternatives can in such cases be altogether avoided by retaining in the original theory only the observation conditionals. The revised theory would be preferable on account of its simplicity. Such a revised theory would then only admit empirically equivalent rivals in a trivial and uninteresting sense. All non-trivial alternatives would posit theoretical entities and principles, and would be less preferable because of the unneeded complexity. The scientist can always, insofar as theory construction goes, opt for simplicity. Likewise, underdetermination would fail for theories that entail an infinite number of observation conditionals without thereby also entailing theoretical entities and principles. In this case, any proposed empirically equivalent rival would have to posit theoretical entities and principles, and thus, again, be less simple than the original theory.

Quine thus explicitly allows for exceptions to the thesis of underdetermination, contrary to what some commentators have claimed.[59] On Quine's view, underdetermination can only obtain for theories that cannot do away with theoretical entities and principles:

---

[59] Newton-Smith (1978, p. 71), for example, spoke of "Quine's notorious claim that (…) *all* theories are underdetermined"; Laudan (1990, p. 271) wrote that Quine holds that "*for any theory, T, and any body of evidence supporting T, there is at least one rival (i.e. contrary) to T that is as well supported as T*";

> Here, evidently, is the nature of underdetermination. There is some infinite lot of observation conditionals that we want to capture in a finite formulation. Because of the complexity of the assortment, we cannot produce a finite formulation that would be equivalent merely to their infinite conjunction. Any finite formulation that will imply them is going to have to imply also some trumped-up matter, or stuffing, whose only service is to round out the formulation. There is some freedom of choice of stuffing, and such is underdetermination. (EESW, p. 324)

Quine makes this same point also in a slightly different context: In a response to Newton-Smith, he says that only "physical theory, the global theory of the world, is underdetermined, but not (…) every subordinate system" (CNS, p. 66). Subordinate systems are partly determined from above, so to speak; they are selected depending how neatly they fit whichever global theory is adopted at the moment. Moreover – and this is where his response to Newton-Smith connects to the passage from EESW just quoted – unmanageably large formulations might be needed to avoid positing theoretical entities and principles. Although a global theory of the world may be expected to entail an infinite number of observation conditionals only by positing theoretical principles and entities, the subsystems that make up the global theory may be simple enough to avoid them altogether.[60]

---

and Devitt (2002, p. 45) suggests that for Quine *any* theory that posits unobservables would have rivals that are empirically equivalent to it.

[60] In the 1980s, Quine replaced the notion of 'observation conditional' with that of 'observation categorical' (see footnote 58 above). Unlike the former, which are "pegged" to individual space-time points and therefore infinite in number, the latter express generalities and are thus finite in number, since only a finite number of generalities can count as observational. It is not clear whether Quine came to allow for the exceptions we have been discussing once the notion of 'observation conditional' is dropped. Nevertheless, it seems to remain true that some theories imply (a finite number of) observational *categoricals* without thereby positing theoretical entities and principles, whereas more complex theories imply observation categoricals which are (also finite in number but) so varied in kind and so numerous that those theories become unmanageably complicated if some theoretical entities and principles are not also posited. It seems reasonable to assume that Quine's view in the 1980s and 1990s would be that only theories of the latter kind can be underdetermined. The difference is then not between the finite or infinite number of observation *conditionals* that a theory implies, but rather between a manageable and an unmanageable assortment of implied observation *categoricals*. However, even in Quine's later view, we would still need a theory containing theoretical principles and entities in order to find the out what the finite list of observation categoricals should be. This is an issue which Quine discusses in connection with Craig's theorem. See EESW, pp. 324-326, Craig (1956), English (1973), and List (1999).

Quine's initial writings on underdetermination claim that the thesis is rendered plausible by the fact that "there is a slack" (PR, p. 254) between theories and observations. One way to understand that slack is by analogy to points on a plane and lines connecting them: Imagine a number of points spread out arbitrarily on a plane. Then, various (in fact, infinite) lines can connect them. Analogously, if we think of the points as observable events, and the lines connecting them as theories or theoretical hypotheses, it seems plausible to think that several alternative theories can be designed to conform to all observations, even all possible observations.[61] Points on a plane do not determine a unique line connecting them; by analogy, observations do not determine a unique theory that predicts them.

As we shall see, however, when we move from points on a plane to actual scientific theories, the analogy becomes rather tenuous. Quine is nonetheless confident enough to assert the thesis, and in part that confidence stems from the support which holism (or, the "Duhem-Quine thesis") lends to underdetermination. Unlike underdetermination, holism is a very basic doctrine in Quine's thought. It connects systematically to almost all aspects of his philosophy, in particular, to his qualms about the analytic/synthetic distinction, to the theses of indeterminacy of translation and ontological relativity, and to his general approach to epistemology, empiricism, and naturalism. An indication of how fundamental he took holism to be is betrayed by the fact

---

[61] Quine speaks of rival theories conforming to "all possible" data or observations in RC (1969, p. 302), RIT (1970, p. 179), EESW (1975, p. 313), NNK (1975, pp. 79-80), RA (1984, p. 294), RRG (1986, p. 156), ITA (1987, pp. 9-10), *PT* (1990, p. 97), and SN (1992, p. 9). But when pressed on the issue by Bergström (1990), he replied that "[i]n treating of the underdetermination of theories it is a poor idea to assume compatibility with all possible data, though in an earlier paper I put it that way; for it is both unrealistic and irrelevant. What matters is that the theories be empirically equivalent, that is, that they imply all the same possible data, or, more precisely, all the same observation categoricals." (CB, 1990, p. 53) In later writings, however – perhaps by force of habit – we still find Quine speaking of theories conforming to all possible observations: SN (p. 9), and *PT* (rev. ed., p. 97), both published in 1992.

that he repeatedly speaks of holism as obviously and trivially true. Answering a criticism from Grünbaum, for example, this is what he had to say:

> Your claim that the Duhem-Quine thesis, as you call it, is untenable if taken nontrivially, strikes me as persuasive. (…) For my own part I would say that the thesis as I have used it *is* probably trivial. (CGC, p. 132)[62]

Holism can of course be understood in various ways, not all of which are trivial. As we have seen in Chapter 2, Quine's version says merely that scientific sentences are not each endowed with their own separate empirical content, and therefore that in revising a theory one may choose from several different alternatives, all of which bring about the same desired overall conformation to the given observations reports. Hence, if a theory implies a false prediction, an individual sentence in the theory can only be deemed responsible if various background assumptions are taken for granted. Strictly speaking, various adjustments in distinct parts of the theory and background assumptions may each suffice to avoid the false implication.

The doctrine, Quine wrote, "must command assent", but "with reservations" (EESW, p. 313).[63] One reservation is that observation sentences do have an empirical content of their own, even when considered individually. However, even these sentences are not independent from the theories in which they are embedded, since the terms which they contain also occur in more theoretical statements. Observation sentences may be taken "analytically", word by word, in which case they function just like the more theoretical statements of a theory; and they may also be taken "holophrastically" (*PT*, p. 7), in which case they are thought fit for immediate testing, without the aid of further assumptions or investigation. Taken holophrastically, observation

---

[62] See also EESW, pp. 313-314; TI, pp. 11-12; and *PT*, pp. 15-16.

[63] See also TDR, p. 268-270.

sentences stand out as an exception to an unqualified assertion of holism; taken analytically they do not.

A second reservation to an unqualified assertion of holism has to do with its scope. In TDE Quine wrote that the "unit of empirical significance is the whole of science" (p. 42). In later writings, he regrets such a "needlessly strong statement of holism" (TDR, p. 268). That strong statement may remain true, he wrote, but only in a "legalistic sort of way" (*ibid*.). It remains true that even observation sentences may be revised in light of theoretical adjustments, and it remains true that statements which are much removed from observations, such as the statements of logic and pure mathematics, are not immune to revision: "Revision even of the law of excluded middle has been proposed as a means of simplifying quantum mechanics." (TDE, p. 43) He later stresses, however, that the minimal unit of empirical significance need not be taken to be the whole of science, but "chunks of it" (TDR, p. 268); that is, "a cluster sufficient to imply an observable effect of an observable experimental condition" (*ibid*.) suffices.[64] Each branch, or clusters of neighboring branches, can thus be thought to imply, together with logic and mathematics, its own set of observation categoricals, and hence to have some autonomy relative to other branches:

> Science is neither discontinuous nor monolithic. It is variously jointed, and loose in the joints in varying degrees. In the face of a recalcitrant observation we are free to choose what statements to revise and what ones to hold fast, and these alternatives will disrupt various stretches of scientific theory in various ways, varying in severity. Little is gained by saying that the unit is in principle the whole of science, however defensible this claim may be in a legalistic way. (EESW, pp. 314-315)

[64] Norton (1994), however, still describes the Duhem-Quine thesis as saying that "theories can only confront evidence as whole".

Granted those two qualifications to the thesis of holism – namely, the separable empirical content of observation sentences taken holophrastically and the relative autonomy of each branch of science – the doctrine remains obviously and even trivially true for Quine: In the face of "adverse observations" a scientist can always elect one among various distinct possible revisions to his overall theory, or to his particular branch of science. No individual statement can be thought to be, on its own, responsible for a theory's false implications, unless, of course, that statement is itself an observation sentence or the long conjunctive sentence that states the whole theory. But even in those cases, various distinct revisions of the theory as a whole can be devised that inactivate the false implication, by either preventing the implication from obtaining or by changing the extensions of some of the terms that occur in it.

Holism thus suggests underdetermination:

> If in the face of adverse observations we are free always to choose among various adequate modifications of our theory, then presumably all possible observations are insufficient to determine theory uniquely. (EESW, p. 313)

The conclusion of the reasoning is here offered as something plausible, yet nonetheless conjectural: Holism "lends credence" without actually establishing underdetermination in any way. The reason why it does not is twofold, as we have already seen in Chapter 2. The first is that although "we are always free to choose among various adequate modifications of our theory", we cannot assume that those modifications will all yield theories that are empirically equivalent. Holism says merely that various modifications may "inactivate" a false implication, but does not say whether the theories that result from those modifications are empirically equivalent. Indeed, it is plausible to think that in most of the relevant cases, alternative modifications will yield theories that are not empirically equivalent. Those modifications would

each "inactivate" the false implication in a different way, and in doing so each may "inactivate" (or "activate") some further implications which remain untouched in alternative modifications.

The second reason is that even if all possible revisions to a given theory yielded only theories which are empirically equivalent, those revised theories might not be sufficiently different to count as rivals, or even alternatives to one another in any nontrivial sense. This seems especially likely if the theories in question are global, and not merely local theories. In the case of global theories, it may just happen that all possible revisions that inactivate a false implication yield theories which are not only empirically equivalent but also theoretically identical, or logically equivalent. They would have the same theoretical structure and would posit the same principles and entities. Whatever differences they might have would then be merely terminological; all sentences of one theory would be reducible via translation to sentences of the other. But theories that are thus intertranslatable are best seen as versions of the same theory, rather than rivals or alternatives. A Chinese physics manual and its Japanese translation are not what we would call "alternative" or "rival" theories.

Additional reasons why holism cannot entail underdetermination can also be offered. Holism can be accounted evident not just on the basis of an abstract reflection on the relation between theories and observations in general, but also because of sufficiently clear historical cases of theories that were reformed in various ways so as to avoid known counter-examples.[65] No such direct evidence, however, can be mustered in favor of underdetermination. Although there are some clear enough cases of alternative hypotheses posited as explanations for the same observations, no clear historical cases of empirically equivalent yet rival theories are available.

---

[65] For a couple of historical cases, see Duhem (1914, pp. 184-188). Lakatos (1976) discusses examples from the history of mathematics.

Known cases of rival theories have turned out to be either cases where there is no empirical equivalence (various versions of competing Ptolemaic and Copernican astronomies, for example), or found to be theoretically identical theories couched in different languages (Heisenberg's and Schrödinger's formulations of quantum mechanics, for example).[66]

Furthermore, there are no clear historical cases of global theories. Hence, there is no clear evidence that global theories, if we ever devise one, will admit empirically equivalent rivals in a nontrivial sense. Some have argued that there is evidence for the underdetermination of local theories, which is a stronger thesis than the one defended by Quine. Understood as a universal and unqualified claim, it is clearly false: Not all local theories have the theoretical complexity that it takes for underdetermination to obtain; and most rival local theories may be inconsistent with the overall global theory adopted at the moment. This, of course, does not mean that local underdetermination cannot obtain for some theories. But it does show that we can only be sure that it has obtained by simultaneously establishing that global underdetermination holds.

To establish the underdetermination of global theories, we cannot rely on the kind of evidence from the history of science often drawn in favor of holism. Unlike holism, this is not a thesis for which there is direct evidence.[67] In fact, the majority of the known historical cases of

---

[66] See Jarrett and Leplin (1991), Kitcher (1993, chapter 7), Hoefer and Rosenberg (1994), Sklar (2000), and especially Massimi (2004) for further examples and discussion. The common perception that early formulations of quantum mechanics by Heisenberg and Schrödinger were equivalent – see, for example, Hughes (1989, p. 44-45) – has been contested by Muller (1997a and 1997b), however.

[67] Stanford (2001) has an inductive argument over the history of science which seeks to establish 'transient underdetermination', that is, the underdetermination of local theories by *currently* available evidence. Transient underdetermination, however, is a weaker thesis than Quine's. Even if local theories have always been underdetermined by the evidence available at the time they were adopted, it may still be the case that global theories do not admit empirically equivalent rivals in a non-trivial sense.

competing theories are not cases of empirical equivalence.[68] Moreover, there are no cases of global theories in the history of science: Physics is the best candidate so far, since it is the only discipline that seems to aims at "full coverage" (*TT*, p. 98). But even physics is yet to attain full coverage, and significant portions and research fields are yet to be unified. So, in this sense, underdetermination must remain somewhat of a conjecture.

Some authors, however, speak of algorithms for creating new theories out of existing ones: Kukla (1998, chapter 5) and Van Fraassen (1980, chapter 3, §§ 2 and 3) are prominent cases. But most of those algorithms produce cases of what Quine would regard as intertranslatable theories. In some cases, it remains unclear whether the artificially constructed "theories" are actually alternatives to be seriously reckoned with in science. In other cases, the algorithms for creating new theories just *are* the manuals of intertranslation.[69] The resulting theory formulations thus do not qualify as rivals in a non-trivial sense. So one question that remains is that of what cases of underdetermination might look like. The thesis of underdetermination needs to be examined not only in regards to its truth or plausibility, but also meaningfulness. It might, after all, be an empty thesis; there might not be empirically equivalent *rivals* in any relevant sense of the word. The very idea of rivalry might not make sense when applied to empirically equivalent theories.

Quine took up this issue in EESW. We have seen that up until then his discussions of the thesis just say that it is plausible given the holistic nature of scientific statements and the "less-than-rigid" connections that obtain between theory and observations. He insists that empirically

---

[68] Massimi (2004) offers an interesting and rather detailed analysis of one such case, and further references.

[69] This is a topic I shall leave aside here, but see Chapter 5.

equivalent theories may conflict with each other, and that we should expect a conflict of this sort to come up eventually. In a passage from 1970, he characterizes such conflicts in terms of 'logical incompatibility':

> Theory can still vary though all possible observations be fixed. Physical theories can be at odds with each other and yet compatible with all possible data even in the broadest sense. In a word, they can be logically incompatible and empirically equivalent. This is a point on which I expect wide agreement, if only because the observational criteria of theoretical terms are commonly so flexible and fragmentary. (RIT, p. 179)[70]

In earlier works Quine had not used the notion of 'logical incompatibility'; in *WO* (pp. 22, 23), for example, he speaks merely of *alternative* theories. This shift in terminology was not motivated by a substantive change in his views, however, although it was not inconsequential either. The move was just an attempt to be more precise in stating what kind of rivalry empirically equivalent theories can be thought to have.

A more substantive review came with EESW, where Quine added a clause to his characterization of underdetermination stipulating that theories are to be regarded as underdetermined only when empirically equivalent but logically incompatible rivals *cannot be rendered logically equivalent by reconstrual of predicates*. 'Reconstrual of predicates', in this context, is nothing but a case of translation, which Quine uses instead because in EESW he narrows the discussion down to empirically equivalent theories already regimented into his canonical notation. Since his canonical notation has no singular terms (see *WO*, § 38), differences in the vocabulary of rival theories can only occur among predicates. The remainder of the vocabulary (variables, quantifiers, and logical constants) is shared. In a later work, he speaks more generally of translatability (EC, 2nd and subsequent printings, 1987, p. 29).

---

[70] See also RC (p. 302): "The totality of possible observations of nature, made and unmade, is compatible with physical theories that are incompatible with one another."

The characterization of rival but empirically equivalent theories as non-intertranslatable aims at overcoming the following objection: The usual way of finding two theories empirically equivalent is precisely through translation. Very seldom can the empirical consequences of two theories be checked one by one to ensure equivalence. This indeed is only possible when the number of observation categoricals implied by each theory is manageably small, which is typically not the case. Hence, empirical equivalence needs to be established by reducing one theory to the other through translation, therefore establishing not only empirical equivalence but also the *theoretical* or *logical* equivalence of the thesis, or at least of the translated versions of the theories. This procedure rules out all rivalry and incompatibilities between theories; it thus brings up the suspicion that all cases of empirical equivalence might also be cases of logical equivalence.

This last concern seems to have been initially raised by Humphries (1970), who pointed out that cases of logical incompatibility (or theoretical conflict in general) between empirically equivalent theories must be distinguished from cases of mere terminological variations. A theory formulation in which all occurrences of "proton" and "electron" are interchanged is trivially distinct from the original theory formulation in which the terms are not interchanged. The observable predictions that each implies are the same, since neither protons nor electrons are directly observable, and the theoretical structure of both theories is also the same. The sole difference lies in that whatever one theory refers to using "proton", the other does so using "electron", and vice-versa. There is an uninteresting and trivial sense in which they may be taken to be alternative theories; they are, to be sure, empirically equivalent yet logically incompatible. But there is hardly any reason or advantage to construe theories so stringently.

Scientific practice and the ordinary man's view agree on this regard, treating such variations as merely terminological, and irrelevant for theory.[71]

Following Humphries, a more general objection can be raised against the thesis of underdetermination: it could very well be that all alleged logical incompatibilities among empirically equivalent theories are due not to substantial theoretical conflict but to mistranslation. This objection was voiced by Dummett:

> … there could be nothing to prevent our attributing the apparent incompatibility [of two empirically equivalent theories] to equivocation. Indeed, if we can establish the empirical equivalence of the two theories, we must be able to find translations from each to the other. (1981, p. 617n.)

The objection says that *in principle* any two empirically equivalent theories are bound to be intertranslatable, even when we do not know how to translate them. Quine addressed this kind of criticism in EESW and NNK. In the former he distinguishes between theories and theory formulations, and concedes that the same theory may be formulated in terminologically inconsistent ways. The proton/electron example mentioned above is an example. Logical incompatibility in such cases would indeed be a trivial matter of equivocation, for surely what is meant by "electron" in one theory formulation is just what in the other is meant by "proton". The incompatibility can be avoided by translating both theories into the vernacular English of technical physics. Quine's formulation of the thesis of underdetermination from 1975 onwards rules out such cases by stipulating that it can only be confirmed by empirically equivalent theories that cannot be rendered logically equivalent through translation.

---

[71] There is of course an issue here of how one ought to define or construe "theory". Quine does not think this has important consequences for the thesis of underdetermination (see *PT*, p. 96). Some authors, however, have challenged his views: see Bechtel (1980), Bar-On (1986), Gemes (1991), and Yalçin (2001).

Dummett's objection says that *in principle* it is absurd to think that any two empirically equivalent theories are not intertranslatable. He thus rules out underdetermination by definition: in principle it would never obtain. Quine's response to that kind of objection offers instead a more open-minded and empirically oriented approach to the question, one that rules nothing out by definition, even if we cannot quite comprehend what the thesis implies. This is in keeping with Quine's empiricism, and his qualms with a priori reasoning in general:

> Terminology aside, what wants recognizing is that a physical theory radically different from ours, with nothing even recognizably similar to our quantification or objective reference, might still be empirically equivalent to ours, in the sense of predicting the same episodes of sensory bombardment on the strength of the same past episodes. (NNK, p. 80)

In a different context, he makes the same point:

> Might another culture, another species, take a radically different line of scientific development, guided by norms that differ sharply from ours but that are justified by their scientific findings as ours are by ours? And might these people predict as successfully and thrive as well as we? Yes, I think that we must admit this as a possibility in principle; that we must admit it even from the point of view of our own science, which is the only point of view I can offer. I should be surprised to see this possibility realized, but I cannot picture a disproof. (*TT*, p. 181)

Interestingly, Quine's response to the objection says that *in principle* we cannot rule out the possibility of empirically equivalent theories that are not intertranslatable. The objection itself was that in principle underdetermination cannot obtain because the conjecture is unintelligible. Quine's response is that in principle we cannot rule it out, even if we cannot quite understand what it entails; that is the sense of a sentence already quoted at the beginning of this paper: "The doctrine [of underdetermination] is plausible insofar as it is intelligible, but it is less readily intelligible than it may seem." (EESW, p. 313) Although we cannot quite know – for now, at least – what kind of theories, if any, would satisfy the requirements posed by the thesis, the thesis itself remains plausible. We have no reason to rule out the chances of eventually

finding theories which we somehow come to think of as empirically equivalent while systematically failing to find a translation manual which would render both logically equivalent.

Some philosophers have pointed out that unintelligible possibilities should not count as possibilities to begin with. It is indeed true that we can hardly imagine what it would be like to find a square circle in a distant corner of the universe, unless by "square" and "circle" one means something other than square and circle. The conjecture that Quine entertains, however, is not of that sort. Although he concedes that it is not quite clear what it entails or what it would take to find empirically equivalent but untranslatable theories or what they should look like, the conjecture itself is not incoherent. In this regard it is unlike the "square circle" conjecture, which is just false, if not meaningless. The question whether there can exist theories that are empirically equivalent but not intertranslatable thus remains an open question, one to be answered as any other in the natural sciences. Perhaps empirically equivalent rival theories can be devised and we will eventually find some. Perhaps we will never find any due to our lack of persistence or sheer ignorance. Perhaps there are further facts about theories and translations, currently unknown – perhaps even forever unknown – which render the conjecture empty. In this last case, the thesis of underdetermination might be regarded false. Right now, however, it remains a plausible thesis, given the slack between theoretical hypotheses and observations in modern scientific theories and the likelihood of cultural developments divergent to our own or to our current one.

The conjecture that supports the thesis of underdetermination is an empirical conjecture like any other. Later we may find out that it is false, even if nothing suggests it today. We may even think later that it never made much sense, as we do today of some scientific theories of the past. At the fringes of scientific research, there are no definite or sharp boundaries between

questions of truth and questions of meaning. This is part of what Quine's writings on the analytic/synthetic distinction show, and it is easily perceptible in some recent empirical conjectures in theoretical physics. Despite the strong evidence supporting quantum mechanics, for example, the persistent debates on its interpretation indicate that it is far from clear what it amounts to or entails. The lack of understanding about what the world should be like, given the truth of that theory, does not prevent science to progress as usual; likewise with some recent conjectures in string theory. Similarly with underdetermination: it remains unclear what global theories that are empirically equivalent but non-intertranslatable should look like, but as far we can tell they might exist. This is not only supported by a coherent conjecture as it is suggested by the lack of a rigid connection between theoretical hypotheses and observations.

If this is true, however, Quine's thesis of underdetermination is not as strong as one might be inclined to believe. What in Quine's earlier writings may have sounded like a categorical statement, in EESW is not much more than a conjecture. Whether a system of the world can have rivals that are empirically equivalent to it but cannot be rendered logically equivalent to it through some manual of translation, he wrote, is an "open question" (EESW, p. 327). What is not a conjecture, but persists for Quine as the *thesis* of underdetermination – a "last ditch version" – is "merely that our system of the world is bound to have empirically equivalent alternatives which, *if we were to discover them, we would see no way of* reconciling by reconstrual of predicates" (EESW, p. 328 – emphasis added). The thesis is thus "vague and modest", and says merely that empirically equivalent theories may be so radically different that, even if intertranslation is possible, it may forever elude us. This mitigated version of the thesis does not rule out – nor assert – the possibility that there might be ways of rendering equivalent through translation any two empirically equivalent theories. It says merely that there are

empirically equivalent global theories which, if we ever find them, will remain non-intertranslatable, either because of our ignorance or because such translations cannot obtain to begin with.

This formulation of the thesis significantly demotes its theoretical significance, as Quine himself acknowledges. Quine claims that the thesis remains nonetheless "vitally important to one's attitude toward science". The thesis adds to the generally accepted claim that theories cannot be deduced from observations: not only can theories not be deduced from observations, but observations alone warrant more than one theory. Alternative, rival theories can be equally warranted by observations. The significance of the thesis is not primarily theoretical, however, since it retains a somewhat conjectural status. Rather, its significance is practical in that it enjoins us to refrain from dogmatically asserting the theories we adopt as the only possible description of the world.

3.3 **Vacillation**

We have seen that Quine's formulations of the thesis of underdetermination changed a few times over the years. Some of those changes were motivated by shifts in reasoning and perspective, while others were merely attempts at a more precise formulation. In papers published in the 1980s Quine gives up talk of theories as opposed to theory formulations:

> It has been urged by many, including me, that scientific theory is underdetermined by all possible data; in other words, that different theories can be empirically equivalent. But this depends on what verbal formulations to count as formulations of different theories and what ones to count rather as different formulations of one and the same theory; and surely this question is philosophically uninteresting. What we are given to compare are the verbal formulations, variously unlike but empirically equivalent. Whether and when to rate them as formulations of the same theory is an inconsequential question of

> words; we can take the theory formulations and let the theories go. (RA, p. 294 – footnote omitted)[72]

In papers of this period, Quine also avoids speaking of 'logical incompatibility' among empirically equivalent theories. Characterizing the differences between empirically equivalent rival theories in terms of lack of intertranslation rendered the notion of logical incompatibility inconsequential for the thesis. Since theories that are empirically equivalent must agree on all observation categoricals, logical incompatibilities among them must occur only among sentences which hinge on observation only indirectly. Suppose that one such sentence is Fx, and that one theory affirms $(\forall x)Fx$ whereas its rival affirms $\sim(\forall x)Fx$. Systematically replace all occurrences of F in one theory with occurrences of, say, F'. One theory will then affirm $(\forall x)Fx$ while the other will affirm $\sim(\forall x)F'x$, thus bypassing the logical incompatibility. Repeat the procedure for all sentences which are contradicted by the rival theory and the result will be a new theory which is *compatible* with its empirically equivalent rival. This move of course does not eliminate the incompatibility of the two original theories, nor is it meant to, but it does show that the notion of incompatibility is here inconsequential (see CB, p. 53). The new formulation can indeed be regarded as a translation of the one from which it originated into a new vocabulary. Thus, the rival theories are no more incompatible than a physics manual written in Chinese and its Japanese translation. In order for there to be incompatibility, the theories must be formulated in the same language.

This "trivial expedient", as Quine calls it (EC, p. 30; RA, p. 294-295), splits up all theoretical terms on which the logical incompatibility hinges. The split terms will have different spellings in each formulation. Since the terms are theoretical and their use is only indirectly

---

[72] See also *PT*, p. 96.

affected by observations, the procedure will be of no consequence for the empirical content of the theory. Following this strategy throughout, cases of logical incompatibility among empirically equivalent theories are effectively rendered innocuous. However, splitting the terms may yield untranslatable terms and sentences (*PT*, §§ 41-43), aside from the ones that may already be there. The languages in which those theories are couched might already be partially untranslatable regardless of the split terms. Such, then, is the conflict that empirically equivalent theories can be thought to have.

As we have seen, Quine changed his formulation of the thesis of underdetermination at least three times: His first formulations, in PR and *WO*, speak rather loosely of empirically equivalent theories that are *alternatives* to one another. In RA, he is more specific about what the differences between those theories are, and describes them as empirically equivalent but *logically incompatible*. In EESW, he adds to his definition the clause that those theories cannot be rendered equivalent by reconstrual of predicates. Later (EC), he speaks more generally of untranslatability instead of "irreconcilable by reconstrual of predicates". In light of these last two formulations, he drops the notion of logical incompatibility as inconsequential, and construes the rivalry that empirically equivalent theories can be thought to have solely in terms of non-intertranslatability.

The analysis of underdetermination that Quine proposed in EESW remained mostly untouched in his later writings. Thereafter it was always a "modest and vague" version of the thesis to which he subscribed. However, he did continue to revise his views regarding a specific problem raised by the thesis, namely, that of how to account for the *truth* of empirically equivalent but rival theories. His earlier writings on underdetermination (PR, *WO*) say nothing about it. But as soon as he introduced the notion of logical incompatibility into his

characterization of underdetermination (RC, RIT), he was forced to take a stand. If empirically equivalent alternatives to our own theory of the world are logically incompatible with it, then they must be conceived as false, or at least false inasmuch as they contradict our own theory. The reasoning is straightforward: If two theories are logically incompatible, then at most one can be true.

Accounting true our own theory and rejecting the alternatives that conflict with it is what Quine would later call the "sectarian" attitude. It is the only coherent option when rival theories are logically incompatible with our own. As we have seen, however, logical incompatibility can be avoided, and in its absence the opposing "ecumenical" attitude becomes attractive. It views non-conflicting empirically equivalent theories as equally true, and allows for each to be used and adopted consecutively, thus contributing to an "added perspective on nature". The ecumenical attitude is recommended by empiricism and its "reluctance to discriminate invidiously between empirically equivalent and equally economical theories" (*PT*, p. 99). The sectarian attitude, on the other hand, is recommended by naturalism:

> Whatever we affirm, after all, we affirm as a statement within our aggregate theory of nature as we now see it; and to call a statement true is just to reaffirm it. Perhaps it is not true, and perhaps we shall find that out; but in any event there is no extra-theoretic truth, no higher truth than the truth we are claiming or aspiring to as we continue to tinker with our system of the world from within. If ours were one of those two rival best theories that we imagined a moment ago, it would be our place to insist on the truth of our laws and falsity of the other theory where it conflicts. (EESW, p. 327)

Oddly enough, Quine explicitly defended both attitudes in two successive essays published in *TT* (1981). In the first essay, in a related discussion of the thesis indeterminacy of reference, he adopts the sectarian line:

> … it is a confusion to suppose that we can stand aloof and recognize all the alternative ontologies as true in their several ways, all the envisaged worlds as real. It is a confusion of truth with evidential support. Truth is immanent, and

> there is no higher. We must speak from within a theory, albeit any of various. (TTPT, pp. 21-22)

Quine takes the idea of a neutral standpoint or of a first philosophy from which theories of the world can be assessed to be hardly intelligible at all. One cannot judge the truth of a theory without looking at the evidence that supports it, and to do so one must have criteria for assessing the evidence. But deciding what to count as good evidence already requires a theory of the world, however rudimentary. Hence, although hypothesizing and even doubting about the truth of our own theories may be common scientific practice, those hypotheses and doubts must still be put forth from within whichever evolving theory one holds at the moment.

Just a few pages later, in the first printing of *TT*, Quine defends the opposing ecumenical view:

> Still let us suppose that the two formulations are in fact empirically equivalent even though not known to be; and let us suppose further that all of the implied observation categoricals are in fact true, although, again, not known to be. Nothing more, surely, can be required for the truth of either theory formulation. Are they both true? I say yes. (EC, first printing, p. 29)

The ecumenical view faces a challenge whenever underdetermination is characterized in terms of logically incompatible theories, since incompatible theories cannot be simultaneously true. What allowed Quine to maintain the ecumenical stance in EC was his revised thesis of underdetermination, which does not make use of the notion of logical incompatibility:

> Being incompatible, the two theory formulations that we are imagining must evaluate some sentence oppositely. Since they are nevertheless empirically equivalent, that sentence must contain terms that are short on observational criteria. But then we can just as well pick out one of those terms and treat it as if it were two independent words, one in the one theory formulation and another in the other. We can mark this by changing the spelling of the word in one of the two theory formulations. (EC, 1st printing, pp. 29-30)

In RA (1984), again this same "trivial expedient" is mentioned (pp. 294-295), and the logical incompatibility of empirically equivalent theories described as a "red herring" (p. 294).

As in the first printing of EC, the trivial expedient is used to justify adoption of ecumenical attitude:

> We are thus left only with empirically equivalent theory formulations that are logically reconcilable. If we subscribe to one of them as true, we can call them all true and view them as different descriptions of one and the same world. (RA, p. 295)

The difficulty that remains for the ecumenical attitude is the naturalistic considerations mentioned above: empirically equivalent theories of the world that rival our own presumably posit alternative sets of theoretical principles and entities. Hence, to say that those theories can be equally accounted true amounts to either admitting various realities or to restricting the truth of a theory to its observable consequences. Quine accepts neither consequence. He sees himself as a realist, and maintains both that reality is unique and that we have good reasons to believe in the truth of the best theories we have. At one point, he tried reconciling the ecumenical attitude with naturalism and realism by stressing that underdetermination is an epistemological thesis; that is, a thesis about warrantedness and evidence, not truth:

> The truth of physical theory and the reality of microphysical particles, gross bodies, numbers, sets, are not impugned by what I have said of proxy functions *and of wildly deviant but empirically equivalent theory formulations*. Those remarks had to do not with what there is and what is true about the world, but only with the evidence for what there is and what is true about the world. I was showing that scientific discourse radically unlike our own, structurally and ontologically, could claim equal evidence and that we are free to switch. Still we can treat of the world and its objects only within some scientific idiom, this or another; there are others, but none higher. (RA, p. 295 – emphasis added)

Here we see Quine laying side by side the thesis of underdetermination and the thesis of ontological relativity. The former, however, is potentially more hazardous to the realist than ontological variations allowed by proxy functions. In the latter, theories vary in ontology but may be rendered equivalent through translation. In fact, a manual for such translations may just be a list of the proxy functions. In such cases, then, theories would not only be empirically

equivalent, but also structurally and theoretically identical. They can be rendered logically equivalent through translation. They would differ in what each counts as objects, but not in how they are interrelated. But then there is very little difference between them, other than terminological. The objects that there are, Quine insists, are the ones the theories we adopt say there are, insofar as those theories are true, regardless of how the objects are named.

> Reference and ontology recede thus to the status of mere auxiliaries. True sentences, observational and theoretical, are the alpha and omega of the scientific enterprise. They are related by structure, and objects figure as mere nodes of the structure. What particular objects there may be in indifferent to the truth of observation sentences, indifferent to the support they lend to the theoretical sentences, indifferent to the success of the theory in its predictions. (*PT*, 31)

Thus it is unclear in what sense theories that vary only in ontology could count as rivals.[73] Strictly speaking, such theories are empirically indistinguishable. As in the "electron/proton" example mentioned above, one may shift from one formulation to the other without thereby committing oneself to a different picture of the world. The cases suggested by underdetermination, on the other hand, are such that theory formulations cannot be converted one into the other by systematic intertranslation. These cases presumably differ in a more substantial manner than the ones suggested by ontological relativity. One expects empirically equivalent but non-intertranslatable theories to be structurally dissimilar, that is, to differ not only in ontology but also in the laws governing the interactions of whichever objects it posits. Differences of this sort do seem to present conflicting accounts of what the world is like, even if the theories are not logically incompatible.

Quine seems to have acknowledged the potential difficulty and attempted to resolve it in various ways. One solution he entertained in the 1980s was to conjoin the alternative theories

---

[73] In fact, it is not even clear whether the theories produced by Quine's proxy-functions vary in ontology at all, since the complement sets are presumably already posited by in the original theory, at least if it couched in a set-theoretical framework.

into a single tandem theory. Quine later dropped that solution (see TI, pp. 14-15): The tandem theory would have separate and redundant lobes, each containing terms and sentences which would not make sense in its counterpart. These untranslatable expressions might even have been created by the trivial expedient of splitting theoretical terms to avoid logical incompatibilities cited above. One lobe must thus remain partially untranslatable and completely redundant, given the other. The idea was thus altogether rejected:

> It is as if some scientifically undigested terms of metaphysics or religion, say 'essence' or 'grace' or 'Nirvana', were admitted into science along with all their pertinent doctrine, and tolerated on the ground merely that they contravened no observations. It would be an abandonment of the scientist's quest for economy and the empiricist's standard of meaningfulness. (RRG, p. 157)

In later texts (TI, p. 14; *PT*, p. 98), Quine also mentions "entelechy" and "phlogiston" as examples of terms that are meaningless from the standpoint of our theory; he also mentions the notion of "the center of the universe" discussed by Poincaré. Theories that contain such terms, he says, cannot be rendered logically equivalent to our own; some sentences will be shared, most notably the observation categoricals, others will be untranslatable.

Prompted by a comment by Roger Gibson Jr. (1998, first edition 1986, pp. 152-153), who pointed out the inconsistency of the sectarian and ecumenical attitudes in *TT*, Quine immediately switched back to the sectarian view:

> The sectarian position, then, is my newly recovered stance on these precarious slopes. It is called for in that last case, where no way is evident of annexing the rival system of the world without adding new terms. Our own system is true by our lights, and the other does not even make sense in our terms. (RRG, 1986, p. 157)

Accordingly, he replaced the ecumenical paragraphs in EC with the following:

> Still we might succeed somehow in persuading ourselves of the empirical equivalence of the two formulations despite finding no way of intertranslation. Then we should indeed recognize the two as equally well *warranted*. We might even oscillate between them, for the sake of a richer perspective on nature. But

> we should still limit the ascription of truth to whichever theory formulation we are entertaining at the time, for there is no wider frame of reference. (EC, second printing, 1987, p. 29)

This new defense of the sectarian position would again be short lived, however. In TI (1990), prompted by Davidson,[74] Quine switched back to the ecumenical position, this time arguing that rival but empirically equivalent theories, purged of inconsistencies by the trivial expedient of splitting theoretical terms, can be regarded as separately true on their own terms:

> … economy, which counted against the tandem idea, is imperative only as an ideal of theory construction and not of language. Meaningful application of the truth predicate, on the other hand, extends to the whole language and is not limited to any particular theory formulation. Empirically equivalent and logically compatible theories can be accepted as true descriptions of the world even if one of them uses terms irreducibly alien to the other. There is no call to fuse them into a single redundant theory. Our language can embrace the full vocabularies of both theories, and our truth predicate can then apply to each on its separate merits. (TI, p. 14)

The vacillation, however, persisted, and in that same year of 1990, Quine again switched back to the sectarian attitude (*PT*, 99-101). He argued that the ecumenical view defended in TI raises problems concerning the language of science. If two rival theories are to be equally and separately true, regardless of the fact that each contains expressions that are foreign and meaningless in the other, then the language of science would have to contain all the terms of both theories. Furthermore, the variables would have to range over objects posited by both. If pursued systematically, the result would be a third theory containing the two original ones. Accounting those two theories true does not differ in any relevant way, he argues, from holding a third theory that is just the conjunction of the two. But this, again, violates the maxim of economy in theory construction. Since the conjoined theories are empirically equivalent, one must be redundant, given the other:

---

[74] See Davidson (1990).

> What is to be gained [with the ecumenical attitude] is not evident, apart from the satisfaction of conferring the cachet of truth evenhandedly. The sectarian is no less capable than the ecumenist of appreciating the equal evidential claims of the two rival theories of the world. He can still be evenhanded with the cachet of warrantedness, if not of truth. Moreover he is as free as the ecumenist to oscillate between the two theories for the sake of added perspective from which to triangulate on problems. In his sectarian way he does deem the one theory true and the alien terms of the other theory meaningless, but only so long as he is entertaining the one theory rather than the other. He can readily shift the shoe to the other foot. (*PT*, p. 100)

This was Quine's last published opinion on the matter. It came with a twist, however. After years fluctuating between the ecumenical and the sectarian views, this time he clearly refrains from asserting his position in a categorical manner. Instead, he acknowledges his own vacillation and demotes the problem to a matter of words:

> The fantasy of irresolubly rival systems of the world is a thought experiment out beyond where linguistic usage has been crystallized by use. No wonder the cosmic question whether to call two such world systems true should simmer down, bathetically, to a question of words. Hence also, meanwhile, my vacillation. (*PT*, pp. 100-101)

It is not too clear what to make of this passage. It seems that Quine is resorting here to the idea that there being empirically equivalent rivals (i.e. non-intertranslatable) to any global theory is a conjecture, and that the thesis which he is willing to assert says only that even if that conjecture is false, we may find empirically equivalent global theories for which *we* will find no manual of intertranslation. But this is still a conjecture, since it remains unclear whether we will indeed find a global theory which conforms to all observations, and even less clear what a rival empirically equivalent theory would look like. We can entertain the thought, but only in imagination. The details are vague, and we cannot anticipate how we would react or what we would say about the truth of those theories. The details of the case are crucial here, and in the

absence of those, very little can be said.[75] There is no determinate cut off point here, but only a gradation. The "fantasy irresolubly rival systems of the world" invites us to imagine a situation very different from that of current science. Indeed, one in which the terms and criteria currently used would by then surely have been modified, or abandoned.

Nevertheless, using the notions we have now, it would be a question of words whether to *say* that rival theories are true. This of course depends on how we are willing to use the word "true", but is not limited to that word. The interesting point to notice here is that in Quine's own philosophical trajectory, the question *became* a matter of words. In the beginning it was not, since underdetermination was thought by him to comprise logically incompatible theories. Once the notion of logical incompatibility is dropped from the formulation of underdetermination, the ecumenical and sectarian views cease to diverge on whether to account alternative theories false. Instead, the problem becomes that of whether alternative theories can be regarded as *meaningful* from the standpoint of the theory we happen to hold and use at that moment. Meaningfulness, however, *is* a matter of words. It is a matter of finding a place and a role for certain words in our description of the world.

"Phlogiston", to cite a notorious example, has no place in our current physical theory. We can understand how it was used in the past, even if we now reject those uses and the false theories in which they were embedded. But where empirically equivalent alternatives to our own theory of the world are concerned, there can be no evidence against rival theories that is not also evidence against our own theory. Hence, expressions that cannot be incorporated into the language of our own theory cannot be rejected as meaningless merely because they only make

---

[75] On this particular shortcoming of the use of thought experiments in philosophy, see Wilkes (1988), chapter 1.

sense in a theory that we know to be false. Whether to regard those terms as meaningful should hinge primarily on how tolerant one is willing to be on matters of linguistic usage, not on matters of fact.[76] The moral of the story, Quine wrote, is that "there are various defensible ways of conceiving the world" (*PT*, p. 102). These are nonetheless conceptions of the same world: reality outruns theory, or so "one is prepared to believe" (*PT*, p. 101).

As mentioned at the beginning of this chapter, Quine's vacillations on these matters are often seen as a symptom of a deeper tension within his philosophy. Considering how mitigated the thesis of underdetermination had become by the end of EESW, and how his reformulation of the thesis allowed the question about the truth of rival but empirically equivalent systems of the world to simmer down to a matter of words, that tension should now appear to run rather superficially. In any case, it should be clear by now that underdetermination is not a fundamental thesis within Quine's system: given its conjectural status, no other doctrine in his system can depend on it; if it were to fail, no consequences would follow for Quine's other doctrines. It is nonetheless a plausible thesis, even undeniable for anyone willing to acknowledge the "less-than-rigid" connections that obtain between theories and observations.

3.4 **Some misreadings of Quine on underdetermination**

We are now in position to very briefly review a few misrepresentations of Quine's views on underdetermination, and to reflect on why they happen so frequently in the literature. Some of these have implications for the larger debate on underdetermination, as we shall see in Chapters 5 and 6. The following examples are fairly representative and characteristic of contemporary discussions of the topic.

---

[76] Discussion of this issue continues in the next chapter.

We begin with Laudan (1990, pp. 271 ff.) and Kitcher (1993, p. 251), who misdescribe holism as the "Quinean ground" for underdetermination. As we have seen, in EESW Quine does say that holism "lends credence" to underdetermination, but he then immediately goes on to explain why it is insufficient to establish it on any firm ground. The thesis itself, he says, has a meager theoretical basis, and its justification is rather conjectural; Quine's view is not, in any case, that it rests on holism alone. In a somewhat related misreading, Laudan (1990, p. 268), Kitcher (1993, p. 250) and Leplin (1997, p. 210) assign to Quine the stronger version of holism explicitly criticized by Grünbaum.[77] Quine, as we have seen, agrees with those criticisms and never really defended that version of the thesis (CBC, p. 132; TI, pp. 11-12; and *PT*, pp. 15-16).

Hoefer and Rosenberg (1994, p. 606) as well as Kukla (1996, pp. 139-140) maintain that Quine is committed to the view that rival but empirically equivalent systems of the world must be logically incompatible. We have seen that already by the late 1970s Quine had abandoned that characterization of the thesis, favoring instead the view that rival but empirically equivalent theories cannot be rendered logically equivalent through translation.

Stanford (2000, p. S8, footnote 8) suggests that Quine asserts underdetermination without actually arguing for it. As we have seen, however, Quine did argue for the plausibility of the thesis. The conclusion of his argument may not be as categorical as one might have expected, but then the problem lies in the expectations of the reader rather than in the argument itself. The argument does not conclude that the thesis is true, merely that it is plausible. Quine's thesis of underdetermination is in fact somewhat disappointing and anticlimactic from a metaphysical point of view. As mentioned at the above, most of the current literature on underdetermination

---

[77] See also Norton (1994, p. 3).

sticks to the assumption that the thesis carries with it significant consequences for the debates on realism and antirealism in the philosophy of science. That is not the case with Quine's view of underdetermination.

Finally, Newton-Smith (1978, p. 71), Norton (1994, p. 3), and Devitt (2002, p. 45) suggest that Quine believes that *all* theories positing unobservables have rivals which are equally warranted by observations. However, as we have also seen, Quine denies that this is so. (CNS, p. 66; EESW, p. 324).

Perhaps a plausible explanation for these frequent misreadings may lie in the fact that Quine is a systematic philosopher whose views on one topic cannot be insulated from the rest of his philosophy. Different aspects of his thought converge into his views on underdetermination, and only in that context can his claim be fully understood and appreciated. A second reason may lie in the fact that the authors just mentioned – and indeed most others in the recent literature on the topic – view underdetermination as a metaphysically loaded thesis: a thesis which, if true, would entail the falsehood of scientific realism or the truth of anti-realism or agnosticism about theoretical entities and principles. For Quine, the thesis does not have any such consequences, nor could it have. This seems to have rendered his description and defense of the thesis rather opaque and intractable to some. Quine himself describes his most important published discussion of the thesis as an investigation into "its meaning and its limits" (EESW, p. 313); for the most part it is not a defense of a thesis but the pursuit of an initially plausible idea that turns out to be assertible only in a very mitigated form. The bulk of that paper aims at dismissing formulations of the thesis which are too strong and therefore untenable. In the end, the claim is mitigated to a "vague and modest" (p. 327) thesis. Perhaps much too vague and modest to be adequately heard in the debates about realism and anti-realism that we currently witness.

Finally, a third reason for the frequent misreadings may lie in the fact that Quine's formulations of the thesis were revised over the years, and in the fact that he changed his mind several times about how to address some of the issues that motivate contemporary debates. In addition to those changes, which have been up until now only partially documented in the literature, there was also a change of subject in Quine's writings. Up until 1975, when EESW was published, Quine's discussions of underdetermination are mostly concerned with establishing its plausibility. Afterwards, however, he was mostly concerned with the question of how to deal with a problem left over from his formulation of the thesis, namely, that of how to account for the truth of rival theories. In the recent literature, little distinction has been made between those two sets of texts. As a result his discussions of the metaphysical question of the truth of rival theories have been thought to be discussions of the formulation and justification of the thesis, which they are not.

4. UNDERDETERMINATION, INCOMMENSURABILITY, AND TRANSLATION

> "These ambiguities, redundancies and deficiencies remind us of those which Dr. Franz Kuhn attributes to a certain Chinese encyclopedia entitled *The Celestial Emporium of Benevolent Knowledge*. In its remote pages it is written that the animals are divided into
>
> a. belonging to the Emperor
> b. embalmed
> c. trained
> d. hogs
> e. sirens
> f. fabulous
> g. stray dogs
> h. included in this classification
> i. trembling like crazy
> j. innumerable
> k. drawn with a very fine camelhair brush
> l. et cetera
> m. that have just broken a vase
> n. that from a distance look like flies."
>
> (J.L. Borges, El idioma analítico de John Wilkins)

4.1 **Introduction**

The thesis of underdetermination has been controversial in the literature. While some philosophers have set out to prove it,[78] others tend to see it as a plausible conjecture, even if defeasible. The thesis is often associated with the philosophy of Quine, but Quine himself argued that not all theories are underdetermined (CNS, p. 66; EESW, p. 323) and that some tempered versions of the thesis are "moot" (EESW, p. 326). Moreover, the claim that "our system of the world is bound to have empirically equivalent alternatives that are not reconcilable by reconstrual of predicates however devious" is, he wrote, "an open question" (EESW, p. 327). As we saw in Chapter 3, What persisted for him as the *thesis* of underdetermination, and not merely a conjecture, was the "vague and modest" claim that "our system of the world is bound to have empirically equivalent alternatives which, if we were to discover them, *we would see no*

[78] For example, Kukla (1996 and 1998).

way of reconciling by reconstrual of predicates" (*ibid*. – emphasis added). In later writings, Quine speaks more broadly of intertranslation instead of reconstrual of predicates (see EC, p. 29). At least for Quine, then, assent to the thesis depends on how it is formulated; not all versions are tenable.

Discrepancies in the formulation of the thesis among various authors are not hard to find, particularly in the recent literature. Assent and dissent to the thesis varies accordingly, depending on just how strong the formulation is. As an illustration of this pattern, I offer the following three examples: Laudan and Leplin (1991) conceive the thesis as saying that "*every* empirically successful theory has empirically equivalent counterparts" (p. 459 – emphasis added). Not surprisingly, given such a broad formulation, they think the thesis "stands refuted" (p. 466). Hoefer and Rosenberg (1994) restrict underdetermination to global theories – as does Quine – and claim that while no *a priori* argument can rule out the possibility of rival empirically equivalent and empirically adequate systems of the world, there remain "formidable" (p. 605) constraints on the realization of that possibility. Finally, Stanford (2001), following Sklar (1975 and 1981), proposes a milder version of the thesis, dubbed "transient underdetermination", which does not require theories to be empirically equivalent in the sense that they imply the same observation categoricals, but merely that they be equally warranted by *currently* available evidence. Transient underdetermination says that "there might be garden-variety alternative hypotheses, *not yet even imagined or entertained by us*, but nonetheless consistent with or even equally well confirmed by all the *actual* evidence *we happen to have at hand*" (2001, p. S7). Transient underdetermination, Stanford argues, is unlike stronger versions of the thesis in that it is well-supported by an inductive argument over the history of science.

Perhaps this small sample of the literature suffices to show that the proponents of the thesis tend to formulate it in weaker terms than those used by its detractors: the weaker the formulation, the stronger the support. Our purpose in this chapter is not to argue for or against any particular version being the best way to formulate the thesis, but to address some concerns that have been raised about its intelligibility. We shall focus only on the second of the three versions mentioned above. It is an intermediately strong version, proposed rather hesitantly by Hoefer and Rosenberg and less hesitantly by Quine himself.

Stronger versions of the thesis have been discussed in the literature and are in general rejected. One of those stronger versions says that *all* empirically adequate theories are underdetermined, for example. That is clearly a false claim, however, since not all such theories have empirically equivalent rivals in a nontrivial sense. If we admit theories that do not posit theoretical entities and principles, then those theories must only admit empirically equivalent alternatives in a trivial and uninteresting sense. Alternatives, in such cases, can at best be linguistically or terminologically distinct renderings of intertranslatable theories. They might even count as versions of the same theories, in the same sense that a physics manual written in English and its translation to a foreign language might count as versions of the same manual. What matters, of course, is not whether they are the same. In a trivial sense they are not. Nonetheless, the purposes, claims, and correctness of the translated manual relative to the corresponding scientific community can be strictly analogous to the role its English counterpart plays in the English-speaking scientific community. Thus, their differences can be conceived as merely terminological; translation may render both logically equivalent. In the absence of theoretical entities and principles, a straightforward and unambiguous translation is to be expected (see below, pp. 93-95). Hence, such theories can only be cases of empirically

equivalent *rivals* in a trivial and uninteresting sense; in any case, *not* the sense envisioned by proponents of the thesis of underdetermination.

The thesis fails also for local theories, theories embedded in an over-arching theory. Local theories are corroborated not only directly by the confirmed observation categoricals they imply, but also indirectly by the correctness of the more comprehensive theories in which they are embedded. Evidence for the more comprehensive, perhaps global, theory is indirect evidence for the local theory that best conforms to it or is entailed by it. Hence, empirically equivalent local theories may be unequally supported by observations in general. Observations that have no bearing on a particular local theory may nonetheless confirm or refute the more comprehensive theory in which is embedded. Thus, the underdetermination of local theories can only be established relative to the underdetermination of the more comprehensive theories in which they are embedded. Additionally, some local theories might not posit theoretical entities and principles. Thus, they will not be underdetermined because of the reasons outlined in the preceding paragraph.

Weaker versions of the thesis, such as Sklar's and Stanford's, on the other hand, seem trivially true, if one grants holism. Notice, however, that the holism required for establishing weak (or "transient") underdetermination is itself weak and indeed trivially true.[79] Weak underdetermination says that at any given moment of our history the theories we hold have alternatives, perhaps unknown, that are just as warranted by the evidence available at that time. Holism, on its turn, says that in the face of adverse observations, a theory can be modified in a number of ways so as to inactivate false implications, that is, prevent the theory from implying a

[79] See chapter 2, CGC, p. 132, and *PT*, pp. 15-16.

false prediction. Most, if not all, theories we currently hold, however, were each the outcome of revisions and adjustments of prior theories, which were rejected because they implied false observation categoricals.[80] Hence, if holism is true, alternative adjustments and revisions were available and must have been left aside for reasons other than conformity to observations. But then weak underdetermination must be true also: the theories we now hold have alternatives which are just as warranted by the available evidence. To be sure, those alternative theories need not be empirically equivalent, or equally warranted by observations. That is not a requirement of weak underdetermination. All that this weaker thesis asserts is that the currently available evidence does not determine one theory of the world; and this remains true even if all the rival theories that there are fail to be empirically equivalent.[81]

Most often what proponents of underdetermination put forth is a stronger thesis. Typically this is a claim about global theories, not about local theories (as may be the case with weak or transient underdetermination). Thus it is not a claim about all theories, but merely about theories which purport to offer, as Quine says, "full coverage" (GWW, p. 98): a theory of all there is. Underdetermination in this sense says that any global theory admits rivals that are empirically equivalent to it, that is, entail the same set of observation categoricals.[82] It is oftentimes presented as a conjecture, since we do not currently have nor have ever had in the past a single global theory that is empirically adequate.

---

[80] On Quine's notion of an 'observation categorical', see EC, p. 27, *PT*, pp. 9-11, and footnote 58 above.

[81] A lot here turns on the notions of 'empirical equivalence' and 'equally warranted by observations'. See Laudan and Leplin (1991), Earman (1993), and Chapter 5 for further discussion.

[82] There is a discussion in the literature about which theories should count as empirically equivalent: see Laudan and Leplin (1991) and Earman (1993). Here it is important only to notice that empirical equivalence in the sense just mentioned is stronger than equal warrantedness by currently available evidence: If two global theories entail the same observation categoricals, then they are equally warranted by observations no matter what the observations are.

Regardless of exactly how the thesis is argued for, at first glance it seems to make a less than transparent claim, since it seems unclear how two empirically equivalent global theories could rival each other. Our purpose in this chapter is to clarify how that rivalry is to be conceived. The issues we will be addressing here have to do with the intelligibility of the thesis and not so much with its truth, insofar as these questions can be pursued independently of one another. In principle, it could seem that any two empirically equivalent theories should be intertranslatable, as Dummett once proclaimed (1981, p. 616n – see citation on p. 60 above). Surely, the objection goes, any two empirically equivalent theories assert exactly the same but in different languages.

The following passage by Quine seems to lend itself to an interpretation that agrees with Dummett's claim:

> But what if, happily and unbeknownst, we have achieved a theory that is conformable to every possible observation, past and future? In what sense could the world then be said to deviate from what the theory claims? Clearly in none, even if we can somehow make sense of the phrase 'every possible observation'. Our overall scientific theory demands of the world only that it be so structured as to assure the sequences of stimulation that our theory gives us to expect. More concrete demands are empty, what with the freedom of proxy functions. (TTPT, p. 22)

The passage seems to imply that if two theories equally conform to every possible observation, any apparent theoretical divergencies they might have are immaterial: empirically equivalent yet rival theories would not make divergent claims about the world. The passage, however, seems to raise a conflict between Quine's professed realism (TTPT, p. 21) and underdetermination. One expects that empirically equivalent theories might diverge on theoretical entities and principles, if underdetermination is true. Realism enjoins us to read scientific theories literally, and to regard the theoretical entities and principles they posit as existent and true. But then empirically equivalent rival theories could in principle make

divergent claims about the world, even if they agree on all observations. This, however, contradicts the passage above.

One easy way to make the passage quoted above consistent with realism would be to regard all theoretical divergencies between empirically equivalent theories as merely verbal, or terminological. Quine rejects that line of thought, however. If all divergencies among empirically equivalent theories were merely verbal, then they would be intertranslatable versions of the same theory, or as logically equivalent theories. Underdetermination requires that theories diverge in more substantial ways.

Insofar as Quine is willing to talk about meaning at all, it can be said that a fundamental doctrine within his system, holism, assigns the minimal unit of empirical meaning *not* to individual sentences or terms, but to theories taken as wholes, or large portions of them. Consequently, translations between theories must also take not sentences or terms as the unit of translation, but theories or large portions of them:

> [I]f the English sentences of a theory have their meaning only together as a body, then we can justify their translation into Arunta only together as a body. There will be no justification for pairing off the component English sentences with component Arunta sentences, except as these correlations make the translation of the theory as a whole come out right. Any translations of the English sentences into Arunta sentences will be as correct as any other, so long as the net empirical implications of the theory as a whole are preserved in translation. (EN, p. 80)

This would seem to entail that if the net empirical implications of two theories are the same – that is, if they are empirically equivalent – they would presumably count as translations of each other even if individual sentences and terms of one theory have no counterparts in the other. But then any two empirically equivalent theories would be translations of one another. Quine rejects that line of thought also, as we shall see below.

Proponents of underdetermination – Quine in particular – have sought to overcome the above mentioned difficulties by redefining the thesis of underdetermination. Instead of merely saying that empirically equivalent theories may conflict each other, the thesis was reformulated so as to avoid cases were divergencies are eliminable through translation. Underdetermination would then say that there are empirically equivalent global theories which, if we manage to find them, will elude all our attempts at intertranslation, regardless of whether they are in fact intertranslatable.[83] The idea that theoretical principles of empirically equivalent theories may not be intertranslatable, however, is not immediately clear. One difficulty is the one already mentioned above: all empirically equivalent theories might seem to be *in principle* intertranslatable. A second difficulty derives from observations by Davidson and Putnam about the notions of conceptual scheme and incommensurability: the idea of an untranslatable sentence seems incoherent; if we manage to identify them as sentences and describe whatever role they might play in a theory, then we would have *ipso facto* translated them. We would not even be able to recognize a sentence as a sentence without thereby translating it.

We shall address these difficulties one at a time. Section 4.2 briefly recapitulates how Quine formulated the thesis of underdetermination, and argues that his writings already contain a sketch of an answer to the first of the two difficulties mentioned above. Section 4.3 reviews some of the ways in which the claim that some sentences are untranslatable has been defended in the literature, and presents those defenses as answers to the second difficulty mentioned above. The discussion will be centered on the so-called thesis of incommensurability, which can and has been formulated in terms of (partial) untranslatability. The examples offered by the defendants of the thesis of incommensurability contain indications of the kind of rivalry that empirically

---

[83] See EESW, p. 327, and Chapter 3.

equivalent theories can be thought to have. Section 4.4 concludes the chapter with a brief discussion of the thesis of indeterminacy of translation. Some authors have maintained that it is inconsistent with the thesis of underdetermination; I argue that it neither affects what is here said about underdetermination and untranslatability nor is inconsistent with Quine's formulation of those theses.

4.2 **Underdetermination and partial untranslatability**

As we have seen in chapter 3, Quine revised his formulation of the thesis of underdetermination over the years. His initial formulations (PR, *WO*) merely say that given the lack of a tight fit between theories and observations, one can expect various *alternative* theories of the world to imply the same observation categoricals, no matter how large the number of observation reports is. Whether theories conform to all observations, or all possible observations, is inessential to the thesis; what matters is that they entail the same observation categoricals.[84] Whatever evidence can be summoned in favor of one theory will then likewise corroborate its empirically equivalent alternatives.

In later writings (for example, RC, RIT) Quine described underdetermined theories as empirically equivalent and *logically incompatible*. This was an attempt to characterize more precisely what kind of rivalry empirically equivalent theories can be thought to have, but it proved to be an inconsequential attempt. All incompatibilities can be avoided by the "trivial expedient" of changing the spelling of theoretical terms which occurs in the incompatible sentences.[85]

---

[84] See CB, p. 53.

[85] See *PT*, pp. 97-98, and pp. 65-66 above.

Quine's final formulation of the thesis merely states that there are global theories which we might somehow manage to recognize as empirically equivalent and yet defy all attempts at intertranslation (EESW, EC, *PT*), that is, we will fail at rendering them logically equivalent by translating both to a common language. Quine's final, "vague and modest" thesis of underdetermination does not rule out the possibility that *all* empirically equivalent theories might be intertranslatable versions of each other. Rather, it states merely that we may find empirically equivalent systems of the world for which we will find no intertranslation. This, he maintains, is a plausible thesis, supported by the "less-than-rigid" connections that obtain between theoretical hypotheses and observations and by the possibility of scientific developments divergent with our own. Whether or not such possibilities will ever be realized is then an empirical question not to be answered by a definition, even if one cannot be completely clear on what they might entail or require.[86]

Still, Quine's characterization of the thesis leaves unclear what such alternative scientific developments might look like, and whether we would recognize them as such if we were to encounter whatever theories they might have produced. To be sure, it can hardly be denied that alternative scientific developments are possible. But that alone provides little clue as to what they might entail and how we would understand them, given that they would be partly untranslatable into the language of our own theories. One way of overcoming the difficulty is to provide an illustration of theories that are empirically equivalent but cannot be rendered logically equivalent through translation. Such an illustration, however, would at most be an imperfect case of underdetermination, since a full-fledged case would have to comprise rival systems of the world yet to be devised. Although physics is a discipline which seeks "full coverage" and may

---

[86] See *TT*, p 181.

one day instantiate the idea of a system of the world, as things stand today physics itself is yet to be unified (quantum mechanics and general relativity are still at odds with each other, for example) and the coverage it provides is less than total. Nevertheless, a partial illustration of the thesis would suffice both to advance our understanding of what it entails and to undermine the objection that all empirically equivalent theories are *ipso facto* intertranslatable.

Quine was not too prodigal when it came to providing examples of what he meant by empirically equivalent but non-intertranslatable theories. There was one case, however, to which he returned several times over the years. It is due to Poincaré, and part of an argument to the effect that the laws of geometry are not experimentally testable, a thesis to which Quine does not subscribe. Poincaré argues that those laws are imposed by us onto what we perceive:

> Beings whose minds were made as ours, and with senses like ours, but without any preliminary education, might receive from a suitably-chosen external world impressions which would lead them to construct a geometry other than that of Euclid, and to localize the phenomena of this external world in a non-Euclidean space, or even in space of four dimensions. As for us, whose education has been made by our actual world, we should have no difficulty in referring phenomena to our Euclidean space. (1952 [1905], p. 51)

Poincaré illustrates that possibility with the following thought-experiment:

> Suppose (…) a world enclosed in a large sphere and subject to the following laws:—The temperature is not uniform; it is greatest at the centre and gradually decreases as we move towards the circumference of the sphere, where it is absolute zero. The law of this temperature is as follows:—If *R* be the radius of the sphere, and *r* the distance of the point considered from the centre, the absolute temperature will be proportional to $R^2 - r^2$. Further, I shall suppose that in this world all bodies have the same co-efficient of dilatation, so that the linear dilatation of any body is proportional to its absolute temperature. Finally, I shall assume that a body transported from one point to another of different temperature is instantaneously in thermal equilibrium with its new environment. (p. 65)

In Poincaré's example, bodies shrink as they move away from the center and become colder. As they move away from the center, however, the proportion between the size of the object and the distance between it and the edges of that space remains constant. If it's half-way

to the edge, its size will have shrunk accordingly; if it's ¾ of the way to the edge, its size will have shrunk even further, and so forth. As a result, by traveling finite distances (in proportion to the size of its body) in finite amounts of time it will never reach the edges of the sphere. The distance between the position where an object is inside the sphere and the edges of that sphere remains always infinite in proportion to the size of the object. Measured from inside the sphere the distance from any point to the edges is infinite, since all measuring instruments will also shrink proportionally as they approach the edges. At the edge, all objects shrink to a point.

Poincaré's example purports to show that observations alone cannot determine the geometrical laws that are true of the space in which we live. In his thought-experiment, the laws of both Euclidean geometry and Bolyai-Lobachevskian geometry accord with observations; choice between them is thus underdetermined by observations. For Poincaré, this would show that they are purely conventional and describe not the structure of space, but the properties we choose to assign to space. Quine used Poincaré's examples as an illustration of thesis of underdetermination, but rejected his conclusions about the nature of the laws of geometry. True to his holism, Quine regarded the laws of geometry as standing on a par with the more abstract and theoretical laws of physics; they meet the test of observations together with less theoretical portions of science and ordinary assumptions about testing methods, behavior of lab technicians, and so forth.

The example, for Quine, serves as an *illustration* of what rivalry between empirically equivalent theories might look like, but no more than that. Insofar as the theory of general relativity is true, both Euclidean and Bolyai-Lobachevskian geometries are false of our physical space. Observations thus conflict with both those geometries when each is assessed in conjunction with general relativity. An actual case of underdetermation would require two

global theories which conform to all observations. Since Poincaré's examples do not, they cannot provide more than an illustration.

Quine picked up the example in the 1970s but changed his mind a few times about its status over the following couple of decades. At first, he conceived it as a bad illustration of underdetermination, because the two rival geometries could be rendered logically equivalent by reconstrual of predicates, or translation. He thought of the two geometries mentioned by Poincaré as similar to the case of a physics theory in which all occurrences of "proton" have been substituted for "electron", and vice-versa. The latter theory and the original one without the substitutions are of course logically incompatible, since they imply contradictory sentences about electrons and protons. For example, one theory might say that protons are positively charged, whereas the other that they are negatively charged. Nevertheless, those two theories are clearly intertranslatable, and can without difficulty be rendered logically equivalent by translating both into the vernacular English of contemporary physics. The two formulations can thus be viewed as versions of the same theory: what one formulation calls "protons" is just what the other calls "electrons". The difference between them is purely terminological. In the 1975, Quine thought that the Euclidean and Bolyai-Lobachevskian geometries which the example by Poincaré brings up would be likewise intertranslatable, although less easily or straightforwardly so:

> [F]ollowing Poincaré, suppose the two theories alike except that one of them assumes an infinite space while the other has a finite space in which bodies shrink in proportion to their distance from centre. Even here we want to say that the difference is rather terminological than real; and our reason is that we see how to bring the theories into agreement by translation: by reconstruing the English of one of the theories. (NNK, p. 80)[87]

[87] See also EESW, p. 322.

No translation was ever provided, however, and later Quine came to accept Poincaré's thought-experiment as containing a good illustration of underdetermination. The notion of 'the center of space' is ambiguous in Euclidean geometry. In fact in any space that is infinite, any point can count as a center. In Bolyai-Lobachevsky geometry, on the other hand, only one point counts as the center. A complete intertranslation must therefore fail; the two theories cannot be rendered logically equivalent by reconstrual of predicates:

> One may ask what to count as distinct theories and what to count merely as different formulations of a single theory. If two empirically equivalent theory formulations are interconvertible sentence by sentence through reinterpretation of terms, surely we should regard them as formulating a single theory. If we just interchange the words 'proton' and 'electron' in a physics treatise, we already have two formulations that are related in that trivial way. On the other hand Poincaré's two theories are not thus interconvertible, since one of them singles out a central point of space while the other provides for no such singularity. (TI, p. 13)

In *PT*, the same point is repeated. The two geometries, Quine says,

> ... differ (…) more deeply than in the mere choice of words. The theory with the finite space makes crucial use of a theoretical term that admits of no counterpart in the theory with the infinite space – namely, 'center of space'. (p. 97)

So here we would have, on Quine's later view, theories that are empirically equivalent but cannot be rendered logically equivalent through translation Partial intertranslation is of course perfectly possible, and indeed some of the sentences – in particular observation sentences and categoricals – are logically equivalent. The notion of 'center of space' of Poincaré's disk geometry and other notions that are directly connected to it fail to have counterparts in Euclidean geometry. Likewise, the notion of 'finite space' has no counterpart in the Bolyai-Lobachevsky geometry.

In the introductory section above (4.1), we saw that in EN Quine stated that theories must be translated with an eye on preserving the "net empirical implications" of each (p. 80). Of

course, that cannot mean that any two empirically equivalent theories count as translations of one another. It is true that the net empirical implications of the two theories in Poincaré's example are indeed the same. Hence, one might conclude, the two, taken as wholes, are translations of one another, even if particular portions or sentences of one have no counterparts in the other. This, according to Quine, would be "a queer translation" (EN, p. 79). In cases where no translation sentence by sentence is available or possible, he wrote, we "might better speak (…) not of translation but simply of observational evidence for theories" (*ibid*). What we should properly call a translation is something which allows for fluent dialogue and negotiation across languages; hence something that requires sentence by sentence translation, even if not word for word.

Among theories that are empirically equivalent, surely all observation sentences and categoricals must be intertranslatable. Such sentences command immediate assent or dissent on the part of all competent speakers of the language that witness the occasion described; they are the sentences upon which translation manuals begin to be built. They are, to use Quine's words, the "entering wedge in the learning of language" (*PT*, p. 5). If any sentences are intertranslatable, then these are surely among them. Theories are empirically equivalent if they entail the same observation categoricals, or, in case they are formulated in different languages, if they entail intertranslatable observation categoricals. Perhaps some or even most of the more theoretical statements are likewise intertranslatable. However, if the two theories are to be thought of as genuinely distinct, rival theories, then there must be at least one sentence or a set of interconnected sentences which cannot be intertranslated. The two theories must be such that no reconstrual of predicates or translation can render them logically equivalent.

Translatability is a relation between two or more languages. In Poincaré's example, all sentences which are true of the Euclidean sphere and contain the phrase 'the center of space' have no counterpart in Bolyai-Lobachevsky geometry. The notion of a center can be defined for an Euclidean sphere S of radius R as follows: Let S be a non-empty set. A metric on S is a function d: $S^2 \rightarrow \Re$ such that (a) for all $x,y \in S$, $d(x,y) \geq 0$; and $d(x,y) = 0$ iff $x=y$; (b) $d(x,y) = d(y,x)$; and (c) $d(x,y) \leq d(x,z) + d(z,y)$. Then, for every point p in S, there is a least number r such that for all $q \in S$, $d(p,q) \leq r$. Call this the radius of space S (from point p). The point in S whose associated radius is the smallest ($r = R$), is the center of S. If the space is Euclidean and finite, as is the case with Poincaré's sphere, that point is unique.[88] Given this definition, if S is metrically infinite, it follows immediately that no point in S is the center of S. Hence, the definite description "the center of space", which denotes a single point in the spherical space of Poincaré's example, denotes infinitely many points in an infinite space.

The defectiveness of the notion of 'the center of space' is not of the sort one frequently finds in ambiguous ordinary phrases such as "the person I spoke with yesterday" or "bank". In the latter, ambiguity can be avoided by further specification, such as "the person I spoke with yesterday whose name is so-and-so" and "by bank I mean 'financial institution'", and so forth.

---

[88] The uniqueness of the center of a sphere can be proved as follows: suppose, for reductio, that the center is not unique. Then there are two points, p_1 and p_2, whose associated radii, r_1 and r_2, are equal to each other ($r_1 = r_2$) and smaller than that of any other point in S which are not equal to either of them. Since $p_1 \neq p_2$, $d(p_1, p_2) > 0$. Pick any number $a \in \Re$ such that $0 < a < d(p_1, p_2)$. Since $p_2 \in S$, $a < r_1$. So, $0 < r_1 - a < r_1$. So there is a $q \in S$ such that p_2 is between p_1 and q and $d(p_1, q) = r_1 - a$. ['Betweenness' is a relation which satisfies at least the following three axioms: (1) If B is between A and C (written A * B * C), then A, B, C are three distinct points on a line, and also C * B * A. (2) For any two distinct points A, B, there exists a point C such that A * B * C. (3) Given three distinct points on a line, one and only one of them is between the other two. (See Robin Hartshorne (2000), pp. 73 ff.)] Since $r_1 = r_2$, $r_1 - a = r_2 - a$. Since p_2 is between p_1 and q, $d(p_2, q) = r_2 - a = r_1 - a$. Given that p_2 is between p_1 and q, it also follows that $d(p_1, q) = d(p_1, p_2) + d(p_2, q)$. So $d(p_1, q) = d(p_1, p_2) + r_1 - a$. Since a was chosen such that $0 < a < d(p_1, p_2)$, it follows that $d(p_1, p_2) - a > 0$. Hence, $d(p_1, p_2) - a + r_1 > r_1$. But $d(p_1, p_2) - a + r_1 = d(p_1, p_2) + r_1 - a = d(p_1, q)$. Hence, $d(p_1, q) > r_1$. So p_1 is not a center of S, because $q \in S$ and $d(p_1, q) > r_1$.

Nothing of this sort can prevent "the center of space" in an infinite space from being non-unique: it is a theorem of Euclidean geometry that (infinite) space has no center. Hence, however the notion is translated into Bolyai-Lobachevsky geometry, the uniqueness condition of the definite description will not be satisfied. Even when the two theories of Poincaré's example are formulated in English, intertranslation fails in the sense that the two theories cannot be rendered logically equivalent. The roles played by the notion of 'center of space' in the two theories are not analogous. In fact, one might even say that it plays no role at all in Euclidean geometry. At best it can be used in theorems stating that there is no center of space, and other theorems which entail its nonexistence. The notion of a center can of course be translated, and may even be understood in the exact same way in both theories. But in one case it is ambiguous, whereas in the other it is not. Hence, the two theories, despite their empirical equivalence, remain rivals.

Similar examples of empirically equivalent but partially non-intertranslatable theories are not hard to devise. Van Fraassen (1980, pp. 49-50), for example, suggests a Newtonian theory according to which the universe rests motionless in absolute space, and a modified Newtonian theory according to which the whole universe moves in absolute space at velocity $w > 0$. His example, however, is not wholly satisfactory: the two theories clearly and trivially can be rendered equivalent through translation, or reconstrual of predicates. What one theory calls "absolute velocity = 0" is just what the other calls "absolute velocity = w".

We can, nevertheless, slightly modify van Fraassen's example so that it can work as an illustration of empirically equivalent theories that cannot be rendered logically equivalent through translation. In a Leibnizian physics, the notions of absolute motion and absolute are eliminated; all motion is relative to some particular set of objects or other. The notion of absolute velocity of Newtonian physics must therefore remain untranslatable into the language of

a Leibnizian physics in which motion is by definition a relational notion. Unlike the case of 'the center of space' discussed above, the translation of 'absolute motion' into the Newtonian theory will result not in an ambiguous term, but in a term that does not denote at all. In this respect, it is very much like the notions of 'phlogiston' or 'entelechy' in relation to contemporary chemistry, for example.

Additional examples can also be summoned. Perhaps a particularly interesting one lies in the various interpretations of quantum mechanics and the debates they have spurred, which indicate that notions that have a central role in one side of the debate fail to denote or are ambiguous in the other. Regardless of what those debates entail, the point to be made here is just that some fundamental notions elude all attempts at intertranslation in the sense that the sentences in which they occur cannot be systematically rendered logically equivalent: there are always some sentences containing those notions which are true in one theory or interpretation and false in another.

Rivalry among empirically equivalent theories can thus be thought to be in principle possible, so long as intertranslation fails to render them logically equivalent. Quine's formulation of underdetermination thus already contains a sketch of what rivalry among empirical equivalent theories should look like. As we shall see below, divergent scientific and cultural developments have brought about theories and even ordinary-language phrases and terms which resemble what Quine had in mind in they, too, cannot be rendered logically equivalent through translation. There is no reason to think something similar could not happen at the level of global theories; hence, the historical examples may aid the understanding of the thesis of underdetermination and offer evidence against the objection that all empirically equivalent theories should in principle be logically equivalent under some manual of

intertranslation. As stated above, the examples mentioned here are at best illustrations, since the theories are either fictitious or not examples of global theories. Nonetheless, it is plausible to think that alternative scientific and cultural developments are likely to produce rivalries among global theories, if global theories are to be produced at all.

4.3 **Translation and commensurability**

Cases of untranslatability have been much discussed in the recent literature under the rubric of "incommensurability". A thesis of incommensurability was put forth by Kuhn (1970 ′1962]) and Feyerabend (1962), which says that there are theories which fail at attempts of intertranslation: some fundamental terms and phrases of those theories might remain always foreign to its rivals. As such, the thesis of incommensurability is akin to the thesis of underdetermination, which also came to be formulated in terms of translation by Quine. What follows in this section is a comparative analysis of the two theses and of some objections that were brought against them.

The notion of incommensurability was notoriously introduced into the philosophy of science by Kuhn and Feyerabend. They did so in analogy to mathematical incommensurability: "The hypotenuse of an isosceles right triangle is incommensurable with its side or the circumference of a circle with its radius in the sense that there is no unit of length contained without residue an integral number of times in each member of the pair." (Kuhn 2000, p. 35) By analogy, theories are incommensurable if there is some term or phrase of one theory which cannot be translated into the language of the other. Such theories may even share most of their language; incommensurability requires only that some terms and phrases are not shared. The examples from Poincaré discussed above provide an instance.

The notion of incommensurability was widely interpreted in the secondary literature as entailing the incomparability of theories. This was explicitly rejected by Kuhn and Feyerabend. Examples of incommensurability often mentioned by those authors include the notion of 'phlogiston', which has no counterpart in the language of contemporary chemistry. The theories themselves, that is, phlogiston theory and contemporary chemistry can nonetheless be compared on the basis of the predictions of observations they make. The fact that some terms of one theory have no place in the other – that the sentences in which they occur have no logical equivalents in the rival theory – does not imply that the theories as wholes cannot be compared on the basis of the observation categoricals they each implies. Other examples brought up by Kuhn are the Newtonian notions of 'force' and 'mass', which have no counterparts in the language of Aristotelian physics, and the notion of a 'neutrino', which has no counterpart in nineteenth century English

Quine and Kuhn came to characterize the theses of underdetermination and incommensurability in somewhat parallel ways. Both began with vaguer characterizations of the differences rival theories, and gradually shifted towards their being partially non-intertranslatable. As already seen (4.2), Quine began his descriptions of underdetermined theories talking of *alternative* and *conflicting* empirically equivalent theories. Later he described them as logically incompatible. Finally, he settled for partially non-intertranslatable: underdetermined theories admit rivals that are empirically equivalent but cannot be rendered logically equivalent through translation. Kuhn, of course, does not have in mind empirically equivalent theories, and the theories he describes are not global either. Nonetheless, he, too, came to think of fundamental conflicts among theories in terms of untranslatability. Incomensurable theories are rivals in the sense the sentences of one theory containing some

fundamental expressions of that theory cannot be systematically translated into sentences of the other theory without yielding absurdities. Hence, the conflicts between empirically equivalent theories that Quine describes with his thesis of underdetermination can be seen as cases of the type of conflict Kuhn described with his thesis of incommensurability. The conflicts Quine has in mind are limited to empirically equivalent global theories, whereas the conflicts described by Kuhn are not restricted in that way. Not surprisingly, both theses faced similar objections: the conflicts they describe were thought to be unintelligible and the theses thought to be incompatible with the scientific realism defended by their authors.

In Kuhn's first writings on the topic, in particular in The Structure of Scientific Revolutions (1970a), incommensurability is presented as a relation between scientific paradigms. These, he wrote, are "universally recognized scientific achievements that for a time provide model problems and solutions to a community of practitioners" (p. viii). Incommensurability would then occur as a result of paradigm changes, or during those changes. In such periods, the "proponents of competing paradigms are always slightly at cross-purposes. Neither side will grant all the non-empirical assumptions that the other needs in order to make its case." (p. 148) Disagreements emerge not only on particular claims about nature, but also on the standards for assessing conflicting claims and on the definitions of fundamental terms:

> Within the new paradigm, old terms, concepts, and experiments fall into new relationships one with the other. The inevitable result is what we must call, though the term is not quite right, a misunderstanding between the two competing schools. The laymen who scoffed at Einstein's general theory of relativity because space could not be "curved" (…) were not simply wrong or mistaken. Nor were the mathematicians, physicists, and philosophers who tried to develop a Euclidean version of Einstein's theory. What had previously been meant by space was necessarily flat, homogeneous, isotropic, and unaffected by the presence of matter. (1970a, p. 149)

In what may have been an unjustified exaggeration, Kuhn further described those disagreements as differences in world-views: "the proponents of competing paradigms practice their trades in different worlds" (p. 150). This passage seems to suggest not only that incomparability may have been on Kuhn's mind in his early work, but also that he probably revised his thesis in later works. The point on which the passage insists was that fundamental conceptual divergencies affect not only the theoretical claims but also how observable objects are perceived. Even what counts as evidence or data can be affected by which particular paradigm is adopted. The ontology of the proponents of rival theories may differ significantly, even if assent and dissent to observation sentences and categoricals remains largely unscathed by paradigm changes.

Howard Sankey (1993) has identified three phases in the development of Kuhn's thoughts on incommensurability. The first phase is the one just mentioned, which dates back to the first edition of The Structure of Scientific Revolutions (1962). It characterizes incommensurability as (i) a relation between paradigms, (ii) the outcome of changes in conceptual apparatuses which partly affect observations, and (iii) as entailing divergent world-views. The second phase occurred the 1970s, beginning with a Postscript that was added to the second edition (1970a) of Structure. During this period, Kuhn thought that comparisons between competing theories required that both be formulated in a single, neutral language (1970b), which would not obtain in the case of competing paradigms: "There is no neutral language into which both of the theories as well as the relevant data may be translated for purposes of comparison." (1979, p. 204) As a result, no exact translation of some fundamental notions would be possible between the two theories. This should not be taken as entailing incomparability, however:

> Most readers of my text have supposed that when I spoke of theories as incommensurable, I meant that they could not be compared. But

> 'incommensurability' is a term borrowed from mathematics, and it there has no such implication. The hypotenuse of an isosceles triangle is incommensurable with its side, but the two can be compared to any required degree of precision. What is lacking is not comparability, but a unit of length in terms of which both can be measured directly and exactly. In applying the term 'incommensurability' to theories, I had intended only to insist that there was no common language within which both could be fully expressed and which could therefore be used in a point-by-point comparison between them. (1976, p. 189)

Despite Kuhn's disclaimer, passages such as these have raised concerns among commentators. It says that although a comparison is possible, no point-by-point comparison is possible; but unless some clarification is given as to what that means, one is left with the suspicion of incoherence: a comparison is possible and is not possible. Kuhn's claim can, however, be more charitably interpreted as merely saying that because some fundamental terms and phrases of each rival theory cannot be translated into the language of another, there might not be any one particular theoretical sentence that one theory asserts and the other denies; hence, no 'point-by-point' comparison. A point-by-point comparison would require that divergencies be narrowed down to the truth values each theory assigns to particular sentences. Such comparison, however, will not be possible whenever divergencies between theories might hinge on terms and phrases that are not intertranslatable, as may happen when theoretical sentences are considered. Some theoretical sentences which are then true in one theory will fail to have a counterpart in rival theories. Nonetheless, there will be agreement on a good portion of the terms and phrases of each theory, particularly the ones more directly linked to observation. Comparisons of the predictions of observations of each theory are thus not affected by theoretical divergencies which hinge on terms not directly fixed by observations. The notions of 'phlogiston', 'neutrino', and 'center of space' are of that nature. Comparisons of what the theories that employ those notions predict can, however, be stated without those terms. Rather, all that is required are observation sentences and categoricals. Thus, even if not all statements of

each theory lend themselves to sentence to sentence or point-by-point comparison, comparison in a broader sense is still attainable.

Kuhn delivered his views in this intermediate period in a somewhat tortuous manner, perhaps due to the changes it was undergoing. He describes comparison between incommensurable theories as made possible by the identification of the referents of each theory, which presumably should be the same. But he also seems to suggest that what counts as the referents of each theory will vary according to the theoretical assumptions each makes. At one point, he compares his thesis of incommensurability with Quine's thesis of indeterminacy of translation (see his 1970b, pp. 164-166). Clearly, however, the two theses are very different, since Quine's thesis says that for any language various manuals of translation are possible that conform to all observable speech behavior of the speakers, whereas Kuhn's thesis says that some sentences of a theory are not translatable into the language of other theories. Kuhn's is thus a thesis that asserts untranslatability or failure of exact translation of some fundamental notions, not a plurality of alternative translations.[89] The association between incommensurability and indeterminacy of translation was later dismissed by Kuhn himself (see his 1976, p. 189).

Kuhn's mature views state what he then calls 'local incommensurability': failure to translate between localized clusters of interdefined terms. Local incommensurability, he says, is often occasioned by change or discrepancies in taxonomic categories. One of Kuhn's favorite examples of such discrepancies is the case of phlogiston theory. Although we can understand what users of that theory were attempting to say with the word "phlogiston" and describe the way they used it, neither the term itself nor any other can play a similar role in contemporary

[89] Further discussion in Sankey (1991).

chemistry: "phlogiston", we now think, does not denote. We know, for example, that according to phlogiston theory, phlogiston is "given off in combustion; it reduces the elasticity and life-supporting properties of air; and so on" (Kuhn, 1983, p. 43). But this description of how "phlogiston" was used does not allow for a sentence by sentence translation of phlogiston theory into contemporary chemistry. The observation categoricals that phlogiston theory and contemporary chemistry entail are not the same. Yet even if we managed to isolate some observation categoricals that are entailed by both, the theoretical principles that in each theory entails those categoricals would remain non-intertranslatable: no sentence by sentence translation can render them logically equivalent. In any translation there are always sentences which are true in one theory and false in the other. The world, according to those theories, is governed by different laws and parceled up in different ways. Thus, even though all observation sentences and categoricals and even some theoretical statements can be rendered equivalent through translation, others cannot. And in that sense, the two theories remain partially non-intertranslatable.

In essays published in the 1980s about related issues, Putnam and Davidson objected to the characterization of incommensurability in terms of partial untranslatability.[90] They suggested that the idea of an untranslatable theory is incoherent or unintelligible. Putnam, for example, wrote that

> … if Feyerabend (and Kuhn …) were right, then members of other cultures, including seventeenth-century scientists, would be conceptualizable by us only as animals producing responses to stimuli (including noises that curiously resemble English or Italian). To tell us that Galileo had 'incommensurable' notions *and then to go on to describe them at length* is totally incoherent. (1981, pp. 114-115)

[90] A similar argument had already been put forth by Wisdom (1974).

According to Putnam, by describing older theories historians of science would be providing us not only with means of comparing old theories with contemporary ones but also with translations of those older theories into contemporary vocabulary. Thus, to describe what is meant by an untranslatable sentence would be incoherent: in doing so one thereby translates the sentence that is supposed to be untranslatable. Furthermore, Putnam argues, translation allows for comparison. His argument has the structure of a *reductio*: If the theories we held in the past and later rejected were incomparable to the ones which succeeded them, then in the history of science theory change would be an irrational affair. Each new theory would be about a wholly new set of entities, unrelated to the things referred to by the older theories. This, he argues, is implausible, since new theories are presented as better descriptions of the same objects and phenomena described by older theories. What succeeding theories refer to must largely coincide with what older theories refer to, even if they do not use the same nomenclature.[91]

In a related discussion, Davidson says of Kuhn's cases of incommensurability that "examples like these, impressive as they occasionally are, are not so extreme". Instead, Davidson points out that "the changes and the contrasts can be explained and described using the equipment of a single language." (1984, p. 184) In particular, he says, "Kuhn is brilliant at saying what things were like before the revolution using – what else? – our post-revolutionary idiom." (*ibid*) Davidson's overall argument conceives the understanding of an alien point of view as based on the *translation* of sentences which express that point of view into our own language. This is exactly what Kuhn and other historians of science would have been doing all along. But then, Davidson argues, the very idea of an incommensurable or untranslatable sentence would be itself incoherent and unintelligible. If we cannot in principle translate a sentence or expect to see

---

[91] See his 1973, pp. 197 ff.

it translated it in the future, why should we to take it to be a sentence? Why not just dismiss it altogether as gibberish, nonsense?

These criticisms by Putnam and Davidson of the idea of incommensurability are somewhat parallel to Dummett's rejection of Quine's notion of empirically equivalent yet rival theories. Both are criticisms of the intelligibility of the idea of an untranslatable set of sentences. Putnam and Davidson, however, raise more general concerns, which affect not only empirically equivalent theories, but any two theories or languages that are allegedly untranslatable.

Quine's proposed mitigated version of the thesis of underdetermination already suffices to set aside criticisms such as Dummett's, since Quine's mitigated thesis of underdetermination says merely that even if empirically equivalent theories can in fact be rendered logically equivalent through translation, we may nevertheless forever remain ignorant of an appropriate manual of translation. The theories would then remain rivals to us, and present alternative descriptions of the world that are equally warranted by observations. Rivalry would exist insofar as the theories posit sets or theoretical entities or principles which we cannot show to be logically equivalent by any translation manual, however devious. Whether the theoretical entities posited by each theory are actually different could, however, in principle remain forever unknown. The thesis of underdetermination, on Quine's mitigated version, says merely that there are empirically equivalent global theories for which, if we find them, we will fail to render logically equivalent through.

Responses by Kuhn and Feyerabend take on a different strategy. They, of course, cannot concede that the examples they have in mind might in principle be intertranslatable. There is no reason to expect that, since they are not cases of empirically equivalent theories. Instead, they

reject the objections by Putnam and Davidson by arguing that the cases of incommensurability that they have brought up are in fact cases of partial, or local, untranslatability. The strategy pursued by Kuhn and Feyerabend is to challenge the premises of the arguments by Davidson and Putnam.[92]

Putnam and Davidson assume that the thesis of incommensurability says that sentences from a given theory may not in general be translatable into our language. The thesis proposed by Kuhn and Feyerabend, on the other hand, says merely that sentences of a given theory may not be translatable into the language of another theory. Sentences containing "phlogiston", for example, cannot be translated into the language of contemporary chemistry. "Dephlogisticated air" may in some cases be explained as denoting, in Phlogiston Theory, the same things that in contemporary chemistry is denoted by "Oxygen" or "oxygen-rich air". However, there is no term or phrase in contemporary chemistry which could play the role of a counterpart of "phlogiston". In other words, the term "phlogiston" in Phlogiston theory fails to denote, according to contemporary chemistry. Partial translation is possible, but there is no sentence or set of sentences which can be regarded as playing the same role in both theories. This is what Kuhn called "local incommensurability", the moderate version of the thesis which he came to adopt in his later writings:

> Most of the terms common to the two theories function the same way in both; their meanings, whatever those may be, are preserved; their translation is homophonic. Only for a small subgroup of (usually interdefined) terms and for sentences containing them do problems of translatability arise. (Kuhn, 1983, p. 36)

Feyerabend makes a similar claim: "the conditions of concept formation in one theory forbid the formation of the basic concepts of the other" (1978, p. 69). The thesis therefore does

---

[92] The following few paragraphs draw considerably from Sankey (1990, pp. 1-11).

not exclude the translations that Putnam and Davidson have in mind, which are not translations of the vocabulary and sentences of one theory into the language of another, but rather from the vocabulary and sentences of one theory into the overall language we currently use and any additions that may be deemed appropriate. This is how Putnam explicitly characterizes the thesis of incommensurability:

> The incommensurability thesis is the thesis that terms used in another culture, say, the term 'temperature' as used by a seventeenth-century scientist, cannot be equated in meaning or reference with any terms or expressions *we* possess. (1981, p. 114)

It is also clear that translation into our overall current language is what Davidson has in mind: his discussion is not focused on translation between the languages of any two given theories but on translation into our overall language (see his 1984, pp. 190-191).

Hence, these specific criticisms by Putnam and Davidson are not criticisms of the thesis of incommensurability as it is understood by Kuhn and Feyerabend, but of a stronger thesis. The weaker thesis, dubbed "local incommensurability" by Kuhn, however, remains unaffected by those criticisms. Even if we can fully translate the vocabulary and the sentences of, say Phlogiston theory, it remains true that that vocabulary and the sentences containing it cannot be fully rendered in the vocabulary and sentences of, say, contemporary chemistry. The very notion of 'phlogiston' makes no sense within contemporary chemistry. It is much like metaphysical terms such as 'entelechy' or 'grace'. We can understand how they were used in past theories and to what they intended to refer, but nothing similar or analogous can be stated without loss using only the limited technical vocabulary of contemporary chemistry. Even if Phlogiston theory and contemporary chemistry were empirically equivalent, there is no reason to expect that the sentences of one could be rendered logically equivalent to the sentences of the other. The

theoretical entities and principles of each differ, and it is implausible to think that there is in this case a manual of intertranslation.

Perhaps the following examples may help: imagine a language without any temporal indexicals, say, the language of the angels according to the Chistian theology. There are things we can say using indexicals which cannot be translated into that angelical language. For example, we can say things like "Thank goodness the war is over!" A translation of this sentence into an atemporal language might be attempted with the aid of a fixed time coordinate system: "Thank goodness the war ended at 4:30 P.M. on 2 May 2006". Both the original sentence and the translation might indeed refer to the same moment in time. But the original sentence says something which is inevitably lost in translation: it says that the war is over, that is, that the war is no longer happening. The fact that the war is now *past* was lost in translation. The fixed time-date coordinate does not convey that information. Of course, we can look at our watches and calendars and figure that out. But this is additional information, which we are adding to the translated sentence: "Thank goodness the war ended at 4:30 P.M. on 2 May 2006 and it is now past 4:30 P.M. on 2 May 2006". This addition, however, reintroduces the temporal indexicals which the translation was supposed to eliminate. The distinction between past, present, and future cannot be conveyed by a date-time coordinate system which does not specify the present moment.[93] Losses in translation of this kind are similar to the ones which happen in

[93] Yet another attempt to get rid of the temporal indexical of the original sentence would be to say something like "Thank goodness the war ended at 4:30 P.M. on 2 May 2006 and this sentence is uttered after 4:30 P.M. on 2 May 2006". But again, if this latter sentence contains no temporal indexical, then the verb "is" cannot be understood to be in the present tense. Instead, we should read it as an atemporal "is", and in that case the resulting claim does not say whether the sentence has been uttered, is being uttered, or will be uttered. Hence, it does not say whether the war has ended in the past, is ending at this moment, or will end in the future. Furthermore, the phrase "this sentence is uttered after 4:30 P.M. on 2 May 2006", when uttered, works like any other temporal indexical, since the moment in time it refers to varies according to the context of utterance: to say "this sentence is uttered" while uttering the sentence

the cases discussed by Kuhn and Feyerabend. Information about the referents intended can be conveyed in a language that does not contain the corresponding terms, but translation may leave out some of the information conveyed by the original sentence.

A second assumption of the arguments by Putnam and Davidson that is rejected by Kuhn and Feyerabend says that in order to understand a foreign term or language, or a rival theory, one must be able to translate it into our own language. Davidson, for example, says that

> On the one hand, it is clear that speech requires a multitude of finely discriminated intentions and beliefs. (…) On the other hand, it seems unlikely that we can intelligibly attribute attitudes as complex as these to a speaker unless we can translate his words into ours. There can be no doubt that the relation between being able to translate someone's language and being able to describe his attitudes is very close. (1984, p. 186)[94]

Both Kuhn and Feyerabend reject this assumption. They point out that understanding can also be attained by *learning* the foreign language without translating it: "We can learn a language or culture from scratch, as a child learns them, without detour through our native tongue" (Feyerabend, 1987, p. 76). Becoming bilingual, however, does not entail being able to translate everything one can understand in one language into the other language. Some terms and phrases can and often do remain untranslated. This is something most bilinguals experience to some extent. In the process of learning a new language, some terms and phrases will be learned in the context of a new culture. When that culture comprises practices and habits that have no counterparts in the native culture, the newly learned terms and phrases may also fail to have

---

is equivalent to say "this sentence is being uttered *now*". See Prior (1959), Gale (1962), Schiffer (1978), Perry (1979), Stalnaker (1981), Chisholm (1982), Austin (1990), and Martens (1994) for further discussion.

[94] See also Putnam (1981, pp. 114-115).

counterparts or analogues in the native tongue. Translation in such cases inevitably fails, or is at best partial.[95]

Sometimes, in ordinary speech translation, one opts to enrich the native tongue by adding foreign terms to it. Most if not all modern languages have incorporated foreign terms to their vocabularies. "Hubris", "noir", and "Gemütlichkeit" are examples of foreign terms used in contemporary English. New phrases and terms are often added without clear restrictions on their use. In the context of a theory, however, there are more or less clear limitations on what can be added. Additions must not contradict whatever is implied by the theory, and must conform to the overall maxims of theory construction. Both these reasons preclude the notion of 'phlogiston' and others from being added to the vocabulary of contemporary chemistry. There is in general no reason to add a term which plays no role and which is independent from the rest of the theory.

We can nonetheless understand terms and phrases which are translatable into the language of a particular branch of contemporary science or all branches taken collectively. Understanding can occur by acquiring some knowledge of how it was used in past theories, which roles it played, to what it was intended to refer, etc. Of course, this is part of the process of learning the specific language to which that term belongs. This might be attained with or without translation. We might want to say that from the perspective of our overall language, that is, the technical vocabulary of the sciences plus ordinary speech, terms like "phlogiston" have been translated by historians. Perhaps we might even want to say that "phlogiston" has never ceased to be a term of the overall English language. Or we might want to say that it was lost and later recovered by the work of historians. In the latter case, "phlogiston" would now be like a

---

[95] A similar point was made by Wittgenstein (1958): "… to imagine a language means to imagine a form of life" (§ 19); "…the speaking of language is part of an activity, or of a form of life" (§ 22).

foreign term (say, "hubris") which is added to the native vocabulary in the overall process of translation. Whichever of these alternatives is the case, the thesis of incommensurability or "local incommensurability" still stands: "phlogiston" and sentences from Phlogiston theory which contain it cannot be translated without loss into the technical language of contemporary chemistry.

Kitcher (1978, pp. 534-535) has persuasively argued that the language of contemporary chemistry can be used to identify various referents of the phrase "dephlogisticated air": in some contexts it was used to refer to oxygen while in other contexts it had no referent. This shows that we can describe how the phrase was used and by doing so come to understand it. But it provides no translation for the term into the language of contemporary chemistry. Phlogiston theory posited the existence of a single "principle", phlogiston, which we can now identify as absence of oxygen in the air in some cases and as nothing at all in other cases. Such attempts at translation thus forcibly eliminate the claim made by the Phlogiston theory that there would be a *single* principle in all such instances. That idea is lost in translations of the term into the language of contemporary chemistry. The ontology of Phlogiston theory cannot be mapped onto the ontology of contemporary chemistry.

The knowledge we may acquire of the use of a term such as "phlogiston" or of any other term from theories we no longer adopt is similar to the learning of a culture which is foreign to our own. We gradually become aware of the circumstances in which a term is correctly used in that theory or culture, but may remain incapable of using anything similar in our own theory or culture. We can come to describe how it was or is used but may remain unable or unwilling to use it in describing the world from our own perspective. In finding out how such terms were used we in fact acquire the technical vocabulary of a foreign language. One may, in special

circumstances, come to use those terms in the first person present indicative tense. But insofar as we know that the theories to which they belong are false, such circumstances must remain extraordinary. The terms learned relate to each other in ways that may have no parallel in our native tongue.

Kuhn illustrates the point with ordinary language examples. He distinguishes two kinds of difficulties in the translation of foreign phrases. Some phrases may be ambiguous, and the translator must vary his English rendering of the phrase according to the contexts in which it is originally used. The French word "pompe" is a good example: "In some contexts (typically those involving ceremonies), its English equivalent is 'pomp'; in other contexts (typically hydraulic), its equivalent is 'pump'." (1983, p. 48) In contrast, some words have no clear cut equivalents in English, even when context of use is taken into consideration. In their translations into English, something gets lost. The net of interrelated notions which they allude to has no counterpart in the language into which it is to be translated:

> … contrast the case of '*pompe*' with that of French words like '*esprit*' or '*doux*'/'*douce*'. '*Esprit*' can be replaced, depending on context, by such English terms as 'spirit', 'aptitude', 'mind', 'intelligence', 'judgment', 'wit', or 'attitude'. The latter, an adjective, can be applied, inter alia, to honey ('sweet'), to wool ('soft'), to underseasoned soup ('bland'), to a memory ('tender'), or to a slope or a wind ('gentle'). These are not cases of ambiguity, but of conceptual disparity between French and English. *Esprit* and *doux/douce* are unitary concepts for French speakers, and English speakers as a group possess no equivalents. (*ibid*.)

These examples help us understand why poetry is so difficult to translate. Poetry relies on spontaneous associations and allusions which the mere utterance of a word brings to the native speaker. Those associations and allusions may be wildly different in foreign languages. For that reason, one can almost define poetry as that which is lost in translation. Likewise, humor is frequently lost, and for the same reason.

A similar case, from inter-cultural studies, is brought up by MacIntyre, in a discussion of Confucian and Aristotelian virtue theories:

> …while both the Confucian and the Aristotelian moralist will see and report one and the same person giving freely and liberally to someone else in need, the Confucian may observe an absence of *li*, of that ritual formality which is an essential characteristic of *jen*, a type of absence necessarily invisible to the Aristotelian, who has no words in either Aristotle's Greek or William of Moerbeke's Latin to translate *li*, an expression captured neither by such words as *ethos*, *hostia*, *orgia*, *teletai* used of religious rituals, nor by such words as *ethos*, signifying the customary and habitual, nor by their medieval Latin equivalents. By contrast, the Aristotelian will observe (…) an example of a disposition as evidencing a particular ordering or disorder of the *psyche*, a conformity or lack of it to what is required a citizen of the *polis*, both understood in terms of an ultimate *telos* conceived in a highly specific way, all of which must be invisible to the Confucian who has no words for *psyche* or *polis* either in the ancient Chinese of Confucius or in the later Chinese of Sung Neo-Confucianism. (1991, p. 110)

The point of these examples is that the roles played by particular notions in theories and in cultures may have no counterpart in other theories and cultures. Thus, at best one can describe its role but not adequately translate it and use it within a foreign culture. One can add the foreign word to the native vocabulary. But if it is to retain the same role it had in the foreign culture, it is bound to remain partially untranslatable from the standpoint of our own culture, in the sense that we can describe its use, but not actually use it to describe the world.

These examples from cultural studies are strictly analogous to the ones from the history of science discussed by Kuhn and Feyerabend. A set of interrelated notions that is used in one theory or culture finds no counterpart in another. In each case one can come to understand how they were used and accurately describe their correct uses, but cannot put oneself to use those notions while maintaining one's own theory of the world or without giving up part of one's native culture. These examples also lend additional credibility to Quine's thesis of underdetermination. They provide illustrations of ways in which interrelated sets of sentences from a particular theory or culture may elude all attempts to render logically equivalent to sets of

statements in rival theories or foreign cultures. Additionally, they show how divergent cultural and scientific developments can produce such sentences.

4.4 **Indeterminacy of translation and underdetermination**

In a story by Borges, the philosopher Averroës finds himself at a loss in the translation of a couple of words in Aristotle's Poetics: "Those words were 'tragedy' and 'comedy'. He had come across them years earlier, in the third book of the Rhetoric; no one in all of Islam could hazard a guess as to their meaning." (Averroës' Search, p. 70) What Averroës lacks, in Borges's story at least, is not so much an adequate supply of words but the corresponding culture in which the theatrical practices of tragedy and comedy had their place. In the absence of such contexts, the two words would remain not only untranslatable, but also largely incomprehensible for him. These cases of untranslatability are to be contrasted with the possibilities allowed for by Quine's thesis of indeterminacy of translation.

The thesis of indeterminacy says that "manuals for translating one language into another can be set up in divergent ways, all compatible with the totality of speech dispositions, yet incompatible with one another" (*WO*, p. 27). The cases we have been discussing so far are not cases of indeterminacy of translation, but rather of partial untranslatability. Indeterminacy says that where one translation manual is possible at all, various others, incompatible[96] with it, are also.

---

[96] In later works, Quine explains the sense in which manuals may be considered incompatible: two manuals equally compatible with the behavior of native speakers "might not be usable in alternation, from sentence to sentence, without issuing in incoherent sequences. Or, to put it another way, the English sentences prescribed as translation of a given Jungle sentence by two rival manuals might not be interchangeable in English contexts." (*PT*, p. 48)

Quine distinguishes two theses of indeterminacy: the indeterminacy of sentences (or "holophrastic indeterminacy") and the indeterminacy of reference (or "ontological relativity"). The former, he says, is a conjecture, whereas the latter admits a trivial proof:

> Woods misinterprets me as saying that the indeterminacy of holophrastic meaning is more serious than that of reference. I would say the opposite. My point was just that the former is conjectural and does not follow from the latter, which admits of trivial proof. The essence of the proof is just that *x* is an *F* if and only if the proxy of *x* is the proxy of *F*. (RJW, p. 728)

Both versions of the thesis make an ontological claim. They are not theses about what can or cannot be known by observation of a foreign speech behavior.[97] Rather, the point is precisely that beyond observable speech behavior there is nothing else to be known: "There is nothing in linguistic meaning beyond what is to be gleaned from overt behavior in observable circumstances" (*PT*, p. 38). Hence, in a theory or system of the world, there is no need to posit the existence of meanings.

The correctness of translation is measured by the fluency of communication and negotiation for which it allows. In principle, alternative manuals can work equally well in that regard. They can be incompatible yet correct insofar as they allow for equally fluent communication across languages. The thesis applies not only to the translation of foreign languages but also to translations of the native speech (*PT*, p. 48).[98] What it shows is that ontology is relative to a manual of translation, and that "the notion of propositions as sentence meanings is untenable" (*PT*, p. 102).

---

[97] For an excellent discussion of this point, see Gibson (1998 [1986]).

[98] A common objection to the theses says that if it is true and applies to our home language, then we cannot really know what we mean by what we say. When referring to an object, do we mean that particular object or the complement of that object, that is, the whole universe minus that object? Which of various possible manuals of translation are we using while speaking our home language? Quine dismisses these objections: "Within the home language, reference is best seen (…) as unproblematic but trivial, on a par with Tarski's truth paradigm. Thus 'London' denotes London (whatever that is) and 'rabbit' denotes rabbits (whatever they are)." (RPR, p. 460)

The thesis of underdetermination, on the other hand, makes no claims about ontology; rather, its point is epistemological: alternative descriptions of the world can be equally warranted. Hence, the thesis of indeterminacy would obtain even if the thesis of underdetermination were false. Even if observations *determined* one particular theory of the world, translation would still be indeterminate. Observable speech behavior is compatible with divergent manuals of translations, even if observations determined one particular physical theory of the world. What remains indeterminate are the notions of traditional semantics, in particular the notion of 'meaning'.

> The indeterminacy of translation differs from the underdetermination of science in that there is only the natives' verbal behavior for the manuals of translation to be right or wrong about; no claims are laid regarding hidden neural mechanisms. If translators disagree on the translation of a Jungle sentence but no behavior on the part of the Jungle people could bear on the disagreement, then there is simply no fact of the matter. In the case of natural science, on the other hand, there is a fact of the matter, even if all possible observations are insufficient to reveal it uniquely. The facts of nature outrun our theories as well as all possible observations, whereas the traditional semantics outruns the facts of language. (ITA, pp. 9-10)

Some authors have questioned the compatibility of the thesis of indeterminacy of translation with the thesis of underdetermination.[99] Underdetermination entails rivalry between empirically equivalent theories, but empirically equivalent theories, given indeterminacy of translation, can be rendered compatible. According to Quine's formulation of the thesis of in the late 1960s and early 1970s, underdetermination should in fact entail the possibility of empirically equivalent yet logically *incompatible* theories. That version of the thesis might indeed bring about irresoluble problems for Quine's thesis of indeterminacy of translation, as some have argued.[100]

---

[99] For example, Bechtel (1980), Bar-On (1986), Gemes (1991), Yalçin (2001).

[100] See, for example, Yalçin (2001).

However, when underdetermination is conceived as entailing empirically equivalent but partially untranslatable theory formulations, the inconsistency dissipates altogether. The fact that there is more than one correct manual of translation show that there are more translations than one might have thought; but that does not show that there a translation between *every* pair of empirically equivalent theories. The thesis of indeterminacy and the modest thesis of underdetermination – as Quine understood it from 1975 onwards – make different but compatible claims. The first says that the correctness of translation rests solely on the fluency of communication for which it allows, and not on some ulterior notion of meaning. The thesis of indeterminacy asserts that ontology is relative to a translation manual, and that the notion of propositions as sentence meanings is untenable. Additionally, the thesis thus recommends ontological simplification: the traditional notions of meaning and proposition might be altogether eliminated from a regimented theory of the world. The thesis of underdetermination, on the other hand, makes an epistemological claim: alternative, mutually untranslatable theories of the world can be equally warranted by observations.

5. RECENT DEBATES ON UNDERDETERMINATION

5.1 Introduction

The thesis of underdetermination has received a lot of attention recently, most of which in connection with a debate about the ontological status of the unobservable entities posited by science. Philosophers have been divided on the issue for some time now. Some have argued that the thesis of underdetermination provides a good reason for rejecting, or at least suspending, belief in the existence of those entities. Belief in them would be unwarranted, if underdetermination is true, because for any empirically adequate theory of the world there would be a rival theory which is equally adequate and postulates different entities. And we would be unable to say which of the two is true.[101] Other philosophers have argued that, on the contrary, our current belief in the existence of those entities is sufficiently warranted by observations, even though we may later on find a better theory which postulates different entities. Philosophers who pursue this line of reasoning usually argue that there is something wrong with the arguments supporting the thesis of underdetermination; for example, that there is something confused about the notion that two theories can be empirically equivalent, or that empirical equivalence cannot be used to establish underdetermination. If given enough time and resources, the argument goes, eventually we will find out that one of any two rival theories is to be preferred because of its theoretical virtues (simplicity, fecundity, conservativeness, etc.), if not on empirical grounds alone.[102]

---

[101] See, for example, van Fraassen (1980), Kukla (1998), and Stanford (2001).

[102] See, for example, Boyd (1975), Kitcher (1993), Norton (1993, 1994, 2003a, 2003b, and 2003c), Laudan (1996), Leplin (1997a), and Massimi (2004).

The main goal of this chapter is to assess the state of play of the debates on these issues. As one might expect, Quine's views cannot be aligned with either side of the dispute mentioned above. Rather, his views comprise a third alternative, one which has been somewhat neglected in the recent literature. But we shall leave that aside for now and return to Quine's views in the next chapter.

Scientific realism and antirealism come in many forms, and it is helpful to begin by making some distinctions. This has been already done in the literature: Kukla (1998, chapter 1), for example, distinguishes the epistemic, metaphysical, and semantic claims associated with those two views.[103] In Kukla's classification scientific realists may merely make the "semantic" claim that our best theories of the world are to be read literally, and not taken to be fictions. A "metaphysical" claim says that the theoretical entities and principles posited by science cannot be reduced to statements about observations. Claims about the theoretical entities and principles posited by science are true or false, or approximately true or false, even independently of observations, just like claims about observations can be true or false, or approximately true or false, independently of other claims about observations. Finally, an epistemic claim associated with scientific realism says that we have sufficient evidence to believe that the claims made by our current science are true, or approximately true. Analogously, the same distinctions can be made in regards to antirealism. An epistemic claim associated with antirealism says that our best current theories of the world may not suffice to identify and describe what there is; we ought to suspend judgment (agnosticism) about claims that cannot be directly confirmed by observations. Antirealism can also make the metaphysical claim that all unobservable entities and principles can be reduced to claims about observations. Finally, antirealism can also be taken to say that

[103] See also Horwich (1991), who uses a similar terminology, but makes different distinctions.

the theoretical claims of science are not to be taken literally, they are mere tools or instruments for the prediction of observations.

The issue which has most concerned contemporary philosophers of science is *epistemic* and mostly focused on whether we are warranted in believing in the existence of the alleged unobservable entities posited by our scientific theories, given the currently available evidence.[104] A more general concern has also been raised, about whether we would have reasons for believing in the existence of the unobservable entities posited by a theory of the world which is fully adequate (that is, conforms to all observations). Perhaps one of the reasons why these issues have taken center stage in the recent literature has to with a very forceful and well articulated formulation of those questions by van Fraassen (1980). Van Fraassen himself is agnostic about whether theoretical entities and principles posited by science (both current science and an ideal future science) are real. But he is not agnostic about the *observable* entities described by science, or about the general principles that describe observations. Instead, he is an epistemic realist in regard to those. He is also a semantic realist throughout, in that he does not take the claims made by science as mere fictions or myths, but as attempts at finding out the truth about the world. Science is to be read literally, he argues, even if we cannot know whether some of its claims are true.[105]

In the recent debates on the topic, the most important argument in support of epistemic antirealism about current science and about science in general (even an idealized science) has been the thesis of underdetermination. Van Fraassen, for example, relies on it in chapter 3 of

---

[104] This is not the only issue, however. The debate is sometimes also about the status of observable entities and sense data or about abstract entities (see Field, 1980).

[105] See McMullin (2003) and van Fraassen (2003).

The Image of Science (1980), and explicitly uses a version of the thesis in pp. 46-47.[106] Kukla makes extensive use of it in his *Studies in Scientific Realism* (1998), and Stanford (2001) also puts forth an argument for the thesis. Accordingly, supporters of epistemic realism have criticized the tenability of the thesis of underdetermination. As we shall see, the debate on the topic has evolved greatly in the last few decades. Section two below reviews some of the most influential arguments from underdetermination in support of epistemic antirealism; section three reviews arguments against underdetermination; and section four summarizes a few of the conclusions reached in this debate.

5.2 **Arguments for underdetermination**

The most general form of the argument for underdetermination can be found in Kukla (1998, chapters 5 and 6) and Bergström (1984 and 1993). Roughly the argument says that there always are empirically equivalent rivals to a given theory, and that we can know that that is true because they can in fact be constructed out of the theories we have. Hence, more than one theory is always warranted by observations. Since we cannot know which of them is true, we must at the very least suspend judgment about the truth of those theories, or at least about the claims they make about what is not observable. Presumably those theories will agree on predictions of observations, since they are empirically equivalent, and diverge on the theoretical principles and entities that they posit. Hence, we need not suspend judgment about all the claims those theories make about the observable reality, but only about claims that transcend observations.

---

[106] Although, as Worrall noted (1984, pp. 67-68), the argument from underdetermination is not systematically or fully developed by van Fraassen.

The argument sketched above purports to hold in general. In particular, it can be extended to global theories yet to be developed and to the theories we currently have, which are not global. The latter are either false (and hence unwarranted) or true and underdetermined (hence, again, unwarranted). In either case, we have reason to doubt the claims they make are true, insofar as they transcend observations. Even if their claims are consistent with all observations, including future observations, an alternative theory would be possible that is equally consistent with all observations and nonetheless posits a different set of theoretical principles and entities.

Three basic arguments in support of underdetermination have been put forth in the recent literature. The first can be found in Quine (1975) and also in Hoefer and Rosenberg (1994). It consists in treating the existence of empirically equivalent rival global theories as a possibility in principle, even if we are unlikely to ever find two such theories. Quine further argues, as we have seen (chapters 2 and 3) that the thesis of holism lends credence to that possibility, and that a weak version of the thesis is not merely a possibility or conjecture, but something he is willing to assert as a thesis, namely, that there are (in a platonic sense, so to speak) empirically equivalent global theories, for which we could find no manual of intertranslation even if we found such theories. This latter clause is epistemic, and renders Quine's version of the thesis weaker than the versions currently debated in the literature. In particular, Quine's thesis does not rule out the possibility that the theories we fail to intertranslate can in fact be rendered logically equivalent through a manual of translation yet to be discovered.

Hoefer and Rosenberg (1994, pp. 601-605), on the other hand, stress not what renders underdetermination plausible, but the numerous preconditions that must be met by global theories so that they can be regarded as empirically equivalent and thus support a plausible formulation of the thesis. Some of these preconditions, they argue, are conceptual and

philosophical, while others are factual and contingent. Among them is the need to specify what to count as cases of empirical equivalence: must empirically equivalent theories entail exactly the same observation categoricals, or is it sufficient that they agree only on the observation categoricals which are true in the actual world? Other requirements also mentioned are the need for a "workable distinction between the observable and the unobservable" (p. 603), the need to rule out "bizarre" theories – "for example, almanac-style lists of observation conditionals" (*ibid.*) – and "empirically equivalent theories constructed out of existing (legitimate) theories by cheap tricks" (*ibid*.), such as the electron/proton example discussed in Chapter 4, and also "theories formulated with Goodmanized predicates, instrumentalized theories that make equivalent predictions by simply discarding the unobservable substructure, theories consisting of only the observation conditionals derived from the original, and so on" (p. 604). Finally, Hoefer and Rosenberg point out (p. 605) that underdetermination can only obtain if the world admits of at least two empirically equivalent global theories, which is not something we can ascertain at this point. They conclude that although "[t]here are no a priori arguments extant to show that empirically equivalent, empirically adequate global theories are impossible for our world … the necessary conditions mentioned above for the existence of such rival theories are formidable" (*ibid.*).

The strong thesis discussed by Hoefer and Rosenberg, we should note, differs from the one for which Quine eventually settled. The one Hoefer and Rosenberg discuss is for Quine a conjecture, "an open question" (EESW, p. 327). On this point (and indeed in others), these authors seem to be in complete agreement, although Quine is also willing to assert a weaker thesis. The conjecture of empirically equivalent, empirically adequate rival global theories is a possibility both for Quine and for Hoefer and Rosenberg, and it is suggested by holism, although

holism on its own is insufficient to establish it (Hoefer and Rosenberg, 1994, pp. 594-595; see Chapters 2 and 3 for Quine's views).

A different approach to underdetermination says that there are algorithms for constructing empirically equivalent rival theories out of existing theories. The strategy was illustrated by van Fraassen (1980, pp. 46 ff.), who recalls Newton's theory of mechanics and gravitation, according to which the centre of gravity of the solar system is at absolute rest: TN(0). Van Fraassen contrasts TN(0) with another theory which only differs from it by postulating that the center of gravity of the solar system is at constant absolute velocity *v*: TN(*v*). Then, he argues, since the absolute velocity of an object is not observable, if TN(0) is empirically adequate, so is TN(*v*), for any *v*. However, at most one of those theories can be true, but we cannot know which.

This general strategy has been pursued with greater sophistication by Kukla (1998), who has argued that empirical equivalent alternatives can be constructed from any given, empirically adequate theory. Kukla puts forth various algorithms for constructing new theories; among them, the following: "given any theory T, construct the rival T′ which asserts that the empirical consequences of T are true, but that T itself is false" (p. 59).[107] Then, the two theories are by definition empirically equivalent yet logically incompatible. Kukla also entertains the possibility of a theory T! which says that T holds when we are looking, but is false when we are not.

Hoefer and Rosenberg, as we have seen above, and others – for example, Laudan and Leplin (1991, p. 441) – argue that theories such as T′ and T! are not legitimate. Kukla rejects that claim. He discusses three reasons that have been put forth for dismissing empirically

[107] Kukla attributes this formulation to van Fraassen (1983).

equivalent rival theories constructed by means of algorithms, and objects to all three. The first says that legitimate theories must not be parasitic on other theories. This was explicitly put forth by Leplin and Laudan: "The new T′ is totally parasitic on the explanatory and predictive mechanisms of T" (1993, p. 13). They argue that in order for a theory to be legitimate, it must "posit a physical structure in terms of which an independently circumscribed range of phenomena is explainable and predictable" (*ibid.*).

Kukla concedes that in some cases the empirically equivalent rivals are indeed parasitic, but argues that parasitic theories have in the history of science been accepted as legitimate, and with good reason:

> Consider the following scenario: (1) theory T has been well confirmed, such that its empirical adequacy is widely believed; (2) it's discovered that one of its theoretical principles is inconsistent with an even more firmly believed theory; and (3) no one can think of any way to describe the empirical consequences of T. (1998, p. 69)

In that case, he argues, we may come to believe in the empirical consequences of T while believing that T itself is false. And that is exactly what T′ says. He further argues that that is exactly what Dennett (1973) and others in the cognitive science community have argued in their specific field of research: intentional psychology has had enormous predictive success relative to other theories in that field of research, and it would therefore be foolish not to use it, but its specific ontology is to be rejected, since it conflicts with physicalism.

Kukla also argues that in the case of T! above, the objection of parasitism does not apply, since it is a modified version of T, and not properly parasitic. In fact, he argues, all reference to T can be eliminated: "Whatever detailed story T tells about the world, we can describe T! by telling the same story and adding the proviso that this story is true only when the world is being observed." (p. 71) This seems true, but then it is unclear what distinguishes T! from the

hypotheses of a Cartesian skeptic, in which case Kukla would be confusing a debate about knowledge in general with a more specific debate within the philosophy of science (see below).

A second objection to the algorithms proposed by Kukla is that of superfluity. This is an objection specifically addressed at theories such as T!, and it says that the hypothesis that the world behaves in a different manner when nobody is looking adds nothing to the original theory. In particular, it cannot be positively confirmed, and cannot yield new predictions of observations. Kukla responds to this objection merely by saying that, likewise, there is no reason to hold on to the equally superfluous hypothesis that the laws of the original theory "*continue* to hold when nobody is looking" (p. 71).

A third objection to Kukla's algorithms was put forth by Leplin and Laudan (1993, p. 13). It says that the issue of whether a theory has what it takes to be legitimate should be deferred to the scientific community. The objection then goes on to say that the theories proposed by Kukla have been systematically disregarded by that community. Kukla responds by saying that this is an unclear criterion. Moreover, he claims, sometimes scientists do defend theories that resemble the ones he has in mind (such is the case, he argues, of Dennett and other cognitive scientists).

Kukla has a number of further arguments in support of his algorithms. I shall not discuss any further his reasons because all his responses ignore a more important point, raised by Stanford (2001) and already mentioned above. Stanford, like Kukla, defends underdetermination. But he opposes Kukla's strategy of arguing for it with algorithms. He maintains that at best Kukla's strategy "earns a Pyrrhic victory" (p. S2) against those who oppose underdetermination. Stanford likens Kukla's algorithms to the skeptical hypotheses, which

purport to pose a problem that no humanly attainable criteria can solve. Choosing between the alternatives put forth by Kukla is just like choosing between the hypothesis that all our experiences are illusions (Descartes's evil genius hypothesis) and the hypothesis that they are not. But then, Stanford argues, "if Cartesian fantasies are the only reasons we can give for taking underdetermination seriously, then there simply *is no* distinctive problem of scientific underdetermination to worry about, for the worry *just is* the familiar specter of radical skepticism" (p. S3). Against that kind of reasoning, a different set of considerations can be brought in. But these are not considerations about scientific "theoryhood", but about knowledge in general. The same holds, Stanford argues, for non-algorithmic examples of empirical equivalents such the hypothesis that everything in the universe is constantly shrinking, or increasing in size, without altering the proportions of observable objects, as in the Poincaré example discussed in Chapter 4.

Similarly, Stanford argues, a 'Craigian reduction' of a theory – that is, the string of observation categoricals implied by a theory – is not itself an alternative theory which deserves to be dealt with within science. The problem raised by underdetermination is whether we are warranted in believing anything beyond observation categoricals and observation sentences. Surely, he argues, we "want the *strongest* set of beliefs to which we are entitled by the evidence" (p. S4). A 'Craigian reduction', by contrast, gives us the weakest set of sentences that is warranted by observations. Presumably no one doubts those sentences – at least not because of underdetermination – since empirically equivalent rivals agree on them. The question which underdetermination raises is, rather, whether we are entitled to believe in a stronger theory, that is, a theory that describes not only what is observable, but also what is not and perhaps cannot be observed.

Stanford also argues that algorithms which produce only empirically equivalent *local* theories – such as the case of TN(0) and TN(*v*) mentioned above – also fail to produce a genuine problem of underdetermination. Instead, they merely rephrase the long-standing philosophical problem of confirmation: "if true empirical consequences of a theory are all that matter to its confirmation, then evidence E confirming theory T will equally confirm T + C (where C is any further claim that does not undermine T's implication of E), thus offering spurious confirmation to C itself" (p. S5). Moreover, following Quine (see Chapter 4), it can be argued that TN(0) and TN(*v*) can be rendered logically equivalent through translation. What one theory means by "absolute rest" is exactly what the other means by "absolute velocity *v*". So it is unclear why the two theories should be regarded as alternatives to begin with. These examples, it should be noted, are of local theories only. (The thesis of underdetermination defended by Stanford differs in this respect from the ones by Quine, Hoefer and Rosenberg, and Kukla, as we shall see below).

Kukla has responded to Stanford's criticism by claiming that there are "some new elements in the classical underdetermination argument against scientific realism that don't figure in the Evil Genius argument" (2001, p. 33). To be sure, underdetermination does seem to bring up a problem which remains even when Cartesian skepticism has been rejected, namely, which of two rival but empirically equivalent theories to call true. This seems to remain a problem even when we grant, against the Cartesian skeptic, that most of our observation reports are veridical. Yet Kukla acknowledges that this might not be "newsy enough to escape the charge of recycling an ancient and profitless debate" (*ibid*.).

Kukla also takes issue with Stanford's dismissal of his algorithms as illegitimate theories. He insists that no clear criterion has been offered as to what to count as a legitimate theory. I

shall set this debate aside, and merely call attention to the fact that Kukla and Stanford have two different theses of underdetermination in mind. Whereas Kukla has primarily in mind the underdetermination of global theories, Stanford's thesis (following Sklar) applies only to local theories. Moreover, Kukla is concerned with elaborating a proof that there are empirically equivalent alternatives to any theory. Stanford, on the other hand, is concerned with the actual historical development of science, within which the problem has a stronger grip on our beliefs.

As an alternative to the Kukla-like strategies for arguing for underdetermination, Stanford offers instead a "new induction over the history of science":

> [W]e have, throughout the history of scientific inquiry and in virtually every scientific field, repeatedly occupied an epistemic position in which we could conceive of only one or a few theories that were well-confirmed by the available evidence, while the subsequent history of inquiry has routinely (if not invariably) revealed further, radically distinct alternatives as well-confirmed by the previously available evidence as those we were inclined to accept on the strength of that evidence. (p. S9)

To be sure, in the history of science the alternatives that succeed a theory thought to be false are hardly ever empirically equivalent to it. Nevertheless, Stanford argues, the fact that the new theories are at least as warranted by the old evidence as the old theories were is sufficient to undermine our confidence in the truth of the theories we currently accept. The theories currently accepted, we should inductively conclude, also have rivals which are yet to be discovered but conform equally well, if not better, to the evidence available now. This is a point also made by Sklar (1981): given the limitations of our imagination, it is reasonable to assume that "there are vast numbers of perfectly acceptable scientific theories … we haven't yet brought to mind" including "innumerable alternatives to our best present theories" and probably even some "more plausible than our own theories relative to present observational facts" (pp. 18-19). Stanford accepts Sklar's claims, and his inductive argument over the history of science purports to support

them. The situation described by Sklar has in fact been our epistemic situation throughout most of the history of science, and there is no reason to think it is not our present situation as well, or that it will cease to be in the foreseeable future.

Stanford mentions a number of historical cases in support of his reasoning; among them, the following: "in the historical progression from Aristotelian to Cartesian to Newtonian to contemporary mechanical theories, the evidence available at the time each earlier theory was accepted offered equally strong support to each of the (then-unimagined) later alternatives" (p. S9). Likewise, he argues, is the case of the historical progression from elemental to early corpuscularian chemistry to Stahl's phlogiston theory to Lavoisier's oxygen chemistry to Daltonian atomic and contemporary physical chemistry. So, too, Stanford claims, in the histories of embryology, thermodynamic theories, theories of electricity and electromagnetism, theories of disease, genetics, and so forth.

Because Stanford's argument applies only to local theories, it can at best establish what Sklar calls "transient underdetermination", that is, the underdetermination of local theories by current evidence. This, however, is a different thesis than the one entertained by Kukla, Quine, Hoefer, and Rosenberg. Moreover, Stanford's thesis is not clearly distinguishable from holism. Holism says that various modifications to a given theory can accommodate a counterexample. But presumably the theories we now have were also the outcome of prior modifications and adjustments intended to avoid counterexamples. So we can conclude from holism alone that our current theories have rivals that are just as warranted by the available evidence, which is exactly what transient underdetermination says. Holism does not entail the thesis of underdetermination which Quine, Kukla and others had in mind because it does not guarantee that the alternative modifications that may be proposed to accommodate counterexamples will all yield empirically

equivalent theories that are not intertranslatable. But "transient underdetermination" does not require empirically equivalent rivals. The examples discussed by Stanford are of older theories which are no more warranted by the evidence available at the time than the theories which succeed them are by that same evidence. But the succession of theories in the history of science is hardly ever a succession of empirically equivalent theories.

5.3 **Arguments against underdetermination**

The most influential attack on the thesis of underdetermination in recent times has been that of Laudan and Leplin (1991). The correctness of their arguments is dubious, as we shall see, but it has the merit of bringing to the debate a systematization which was lacking up until then. They purport to attack "both the supposition of empirical equivalence and the inference from it to underdetermination" (1991, p. 449). Their conclusion is not, however, that there can be no cases of empirical equivalence, but merely that we have no guarantee that any given theory has an empirically equivalent rival – as we have seen, this is a point which Quine and Hoefer and Rosenberg grant. Moreover, Laudan and Leplin have argued that some of the arguments that show the empirical equivalence of partial or local theories fail to take properly into account the role of auxiliary hypotheses. In particular, they argue that we can only judge the empirical equivalence of two theories given the state of science and technology at a given time. Future developments in theory and technology may render observable what is presently unobservable, or, to put it in Quinean terms, may enlarge the set of sentences which can count as observation sentences. The empirical content of a theory may change without any changes being made to any of the sentences of the theory itself, if the auxiliary assumptions change. Hence, the claim that two theories are empirically equivalent may always be defeated by future developments. Of

course, by pursuing this line of reasoning, they purposefully disregard empirically equivalent theories created by algorithms, as proposed by Kukla. They argue that those theories are not legitimate alternatives.[108] Instead, they focus primarily on alleged cases of empirical equivalence which have risen more or less naturally in the course of the history of science – an example would be the case of spectroscopic anomalies in the early 1920s discussed by Massimi (2004).

Laudan and Leplin's main argument against categorical assertions of empirical equivalence has three main premises: (1) "[a]ny circumscription of the range of observable phenomena is relative to the state of scientific knowledge and technological resources available for observation and detection" (p. 451), (2) "[t]heoretical hypotheses typically require supplementation by auxiliary or collateral information for the derivation of observable consequences" (p. 452), and (3) "[a]uxiliary information providing premises for the derivation of observational consequences from theory is unstable in two respects: it is defeasible and is augmentable [by developments in theory and in technology]" (*ibid.*). All three premises, they argue, are plausible. Hence, given (1), we cannot anticipate which of a theory's now unobservable consequences may become observable in the future. Furthermore, given (2), a theory's empirical consequence class must be allowed to include statements deducible from the theory only with the help of auxiliaries. And from these and (3), they conclude that apart from shifts in observational status, a theory's empirical consequence class may increase through augmentations to the theory's total consequence class. Therefore, the empirical consequence class of a theory cannot be determined a priori, but is relative to a particular state of scientific practice. Hence, the claim that there are cases of empirically equivalent theories would at best be a claim about the state of science at a given moment. Empirical equivalence, Laudan and Leplin

[108] See also Leplin and Laudan (1993).

argue, is always transient, and defeasible by the future progress of science. Their argument, however, is limited to local theories, which admit indirect evidence from auxiliary assumptions.

Laudan and Leplin also argue (1991, pp. 459-466) that even if there were clear enough cases of empirically equivalent theories, we still cannot conclude that (local) underdetermination is true. Empirical equivalence would not entail (local) underdetermination. They argue that that inference is based on the wrong assumption that theories that have equal empirical consequence classes must have equal evidential support, and point out examples of theories which are supported by evidence not belonging to its empirical consequence class. The gist of the argument is that local theories may be supported by more comprehensive theories, which in turn are independently supported by other evidence.

Kitcher (1993) adopts a similar position, and dismisses (local) underdetermination as the outcome of an oversimplified view of scientific practice. In particular, he argues that the hypothetico-deductive method affords very little clue on these matters. Instead, scientific practice is best understood as conforming to some version of what is sometimes called "eliminative induction", also known as "demonstrative induction". Kitcher argues that because of that oversimplified view of science provided by the hypothetico-deductive method, one may be inclined to believe that all theories that conform to observations should be equally warranted, or that an infinity of alternative hypothesis are just as plausible as the ones that end up adopted by the scientific community. This, he claims, is false: "[a]gainst the background of prior practice there may be only a finite set of possibilities" (p. 247). Most *logically* possible alternatives are just ruled out by theses which are independently supported by the evidence, and by the kinds of research which are allowed the state of technology at a given time. Only a few alternative hypotheses are taken seriously by the researchers, and these are chosen not only on the basis of

the empirical consequences they are thought to bring to a theory, but also on how well they conform to other hypotheses and auxiliary assumptions for which there is considerable independent warrant. Acceptable alternatives are thus constrained by "explanatory dependencies and the views of projectability present in prior practice" (p. 251).

Massimi (2004) follows Laudan, Leplin, and Kitcher on this point, and purports to show how one particular episode in the history of science – that of the spectroscropic anomalies in the 1920s – is best described not as a case of empirical equivalence, which is what the hypothetico-deductive method would lead us to believe.[109] Massimi sees the overarching influence of the hypothetico-deductive picture of scientific method as responsible for the widespread belief in underdetermination. Like Kitcher, Laudan, and Leplin, she defends instead a view of scientific method in which induction plays a more prominent role. The specific view defended by Massimi (2004) is "demonstrative induction". On this view, hypotheses are induced from observation but constrained by already accepted scientific theses and practices in such a way that only one hypothesis actually can accord with the observations. The selection of that particular hypothesis is thus deductive, given the constraints. As in other inductive accounts of scientific method, hypotheses do not "meet the tribunal of experience" *en bloc*, as a simplified version of the hypothetico-deductive method would have us believe.[110] On the contrary, Massimi argues, an analysis of historical cases shows that hypotheses are not warranted exclusively by the observational consequences that they entail. Rather, they are also warranted by indirect evidence from independently well-established theses. She argues that an analysis of the historical scenario

---

[109] Similar arguments, for other branches of science, can be found in Norton (1993 and 2003), Bonk (1997), and Cordero (2001).

[110] For a more detailed description of the various versions of inductive methodology in the sciences, see Norton (2003a and 2003b).

in which the spectoscropic anomalies of the 1920s emerged shows that if we describe the competing explanations along the lines of the hypothetico-deductive method, then it would seem that various empirically equivalent alternatives were put forth (due to Bohr, Heisenberg, and Pauli), and that none of them was better warranted than any other. On Massimi's account, however, only Pauli's theory (the one that eventually prevailed) was warranted, since only that alternative was also supported by all the theoretical assumptions which were independentely warranted at the time (Bohr's Building-up Principle, Bohr-Sommerfeld quantum conditions, Sommerfeld's space quantization, and so on) and a number of "phenomenal premises" which were also acknowledged at the time (X-ray doublets, optical doublets, decrease of the number of states, multiplets, anomalous Zeeman effect, etc.). Pauli's model, she argues, was better warranted than the others available at the time because it was the only one to be supported by all the indirect evidence available in that field of research. The choice for that particular theory was *not* arbitrary or something to be explained exclusively by the sociology of scientific practice.

Similar views about the selection of hypotheses can be found in Friedman (1981) and Norton (1993, 2003a, 2003b, and 2003c). This is a point about which Quine has been targeted (in particular, by the authors discussed in this section). Indeed, he does in various passages explicitly present the relation between theories and observations in terms of the hypothetico-deductive method, and he also connects that characterization with the thesis of underdetermination:

> Scientists invent hypotheses that talk of things beyond the reach of observation. The hypotheses are related to observation only by a kind of one-way implication; namely, the events we observe are what a belief in the hypotheses would have led us to expect. These observable consequences of the hypotheses do not, conversely, imply the hypotheses. Surely there are alternative hypothetical substructures that would surface in the same observable ways. (EESW, p. 313)

The general line of criticism of Quine on this point is that he seems to set aside the intricacies of theory formulation and selection. His views, in other words, do not take into consideration the actual practices of working scientists. This is a criticism which, I guess, Quine would accept. His main concern never was that of describing in any detail the actual practices of scientists. Rather, he mostly addresses a more abstract and general problem: "My general concern has been with the central logical structure of empirical evidence. In fused phrases of Kant and Russell, it is a question of how our knowledge of the external world is possible." (*PT*, p. 18) At that most general level of inquiry, he describes the relation of stimuli and theory in terms of the hypothetico-deductive method. When the analysis of specific cases comes into question, however, Quine seems willing to acknowledge exceptions to his general characterization:

> We must recognize (…) a significant degree of idealization in the foregoing account of hypothesis-testing. The scientist does not tabulate in advance the whole fund of theoretical tenets and technical assumptions, much less the commonsense platitudes and mathematical laws, that are needed in addition to his currently targeted hypothesis in order to imply the observation categorical of his experiment. (*PT*, p. 17)[111]

Quine illustrates the point with a brief discussion of a "near-platitude":

(1) Sodium chloride dissolves in water.

This, he argues, is "[i]n general (…) accepted as a vague statement of strong probability, open to question only where the improbable counter-instance can be plausibly accounted for" (*PT*, p. 18). In most cases, he adds, "near-platitudes" such as (1) and indeed most of science, will remain shielded from empirical disconfirmation, precisely because of the considerations

[111] See also TDR, p. 268, where Quine says of his own description of the hypothetico-deductive method as a "schematic caricature of the experimental method", which "catches the essence of experimental testing" but may need to be qualified in important ways if it is to be seen as a description of actual scientific practice.

listed by Massimi. At a more general level of analysis, however, observation categoricals as well as more general scientific laws remain "responsive somehow to sensory stimulation both early and late" (*ibid.*), "so long and insofar as science has not parted its empirical moorings" (*ibid.*). Science can and occasionally does part those moorings, Quine acknowledges, and when that happens his proposed general schematism does not apply. But even where it does not part, analysis can be conducted at various levels of abstraction, and his own very abstract analysis need not be seen as incompatible with a more specific, historical analysis, such as Massimi's. The point of characterizing the relation of science to stimuli in very general terms is merely to highlight the broadest features of this activity which we call science: "When I cite predictions as the checkpoints of science, I do not see that as normative. I see it as defining a particular language game, in Wittgenstein's phrase: the game of science" (*PT*, p. 20).

Laudan, Leplin, Kitcher, Friedman, Norton, and Massimi surely have a point when they say that an oversimplified description of the methods of science as hypothetico-deductive makes the thesis of underdetermination seem more plausible than it ever is for practicing scientists.[112] It makes a mere logical possibility look like a relevant empirical alternative. Rather, when formulating and choosing new hypotheses and theories, most logical possibilities are disregarded outright, and done so on the basis of accepted practices and theses which are taken to be out of question at that particular moment.[113]

---

[112] See also Hacking (1984), who points out a difference between theoretical scientists and experimental researchers in this regard. The former are much more at ease with alternative hypotheses than the latter. These, while practicing their trade, cannot but see the theoretical entities they use as tools that are ultimately responsible for the production of observable results. Their attitudes towards those tools are thus bound to be realistic in a much stronger way than those of philosophers and theoretical scientists.

[113] This is connected to Friedman's claim (1981) that theoretical postulation, explanation, and confirmation are but aspects of the same general methodological process.

Although this criticism of an oversimplified account of the hypothetico-deductive method seems quite cogent, the same cannot be said of the objections that have been laid against underdetermination. One example is the objection by Laudan and Leplin (1991) and Massimi (2004), for example, which says that empirical equivalence does not entail underdetermination because hypotheses and theories which imply the same observation categoricals can nonetheless be discriminated on the basis of indirect evidence. Two theories, A and B, may be empirically equivalent and yet receive indirect confirmation from another more comprehensive theory, T, which implies, say, A but not B. An obvious shortcoming to this line of reasoning is that it only applies to local theories.[114] At best, the argument shows that the underdetermination of those theories is dependant on the underdetermination of so-called "total theories", that is, theories which incorporate all the auxiliary assumptions warranted at a given time.[115] A total theory may or may not be global, that is, it may or may not imply all the true observation categoricals that there are, or even those that are known to be true; it may, in other words, be a partial theory in the sense that it does not attempt to explain or predict all the phenomena of a certain kind. The objection is then that total theories do not rely on external auxiliary assumptions, because all the warranted auxiliary assumptions are already part of the theory. Indirect evidence can only come from warranted auxiliary assumptions outside the theory, but in the case of total theories, there are none.

This criticism of Laudan and Leplin's argument was somewhat anticipated by Earman (1993), who pointed out that the notion of empirical equivalence can be defined in various ways. If the notion is made strong enough, then surely all empirically equivalent rivals are equally

---

[114] See Kukla (1998, chapters 5 and 6), and Okasha (2002).

[115] See Kukla (1998), and Leplin (1997).

warranted by the evidence. Laudan and Leplin chose an adequately weak notion, from which it does not follow that all empirically equivalent rivals are equally warranted, and then concluded that the argument for underdetermination from empirical equivalence is invalid. Adapting from Earman's paper, we can distinguish three notions of empirical equivalence: Two theories can be empirically equivalent, if (1) no 'crucial test' can set them apart, (2) they imply the same observation categoricals, or (3) they imply the same observation categoricals, and all indirect evidence in favor of (or against) one theory is also evidence in favor (or against) the other. The first is obviously the weakest notion, and the third is the strongest. Laudan and Leplin used the second notion, and for that reason were able to claim that underdetermination does not follow. But if we use the third notion, as Earman proposes, then underdetermination does follow from empirical equivalence. However, empirical equivalence then becomes hard to prove, or argue for – just as hard, in any case, as the stronger versions of the thesis of underdetermination. In the case of total theories, the second and third notions collapse, because there is no indirect evidence to consider. Laudan and Leplin's argument thus relies on an equivocation.

Another problem of Laudan and Leplin's argument is that it uses two principles of confirmation which, if conjoined, yield absurdities. This was pointed out by Okasha (1997). The two principles in question were originally formulated by Hempel (1945), who calls them the "converse consequence condition" and the "special consequence condition". The doctrine Laudan and Leplin aim to criticize is what they call "consequentialism", which is a simplified version of hypothetico-deductivism according to which "hypotheses are to be tested exclusively by an exploration of the truth status of those empirically decidable statements which they entail" (1991, p. 470). As we have seen above, they argue that "being an empirical consequence of an hypothesis is neither necessary nor sufficient for being evidentially relevant" (p. 460). Laudan

and Leplin attempt to show the falsehood of that assumption by discussing a particular historical case, the theory of continental drift (TC), which says that in the past "every region of the earth's surface has occupied both latitudes and longitudes significantly different from those it now occupies" (p. 461). Laudan and Leplin argue that TC is "committed" to the following two hypotheses:

> H_1: There has been significant climatic variation throughout the earth, the current climate of all regions differing from their climates in former times.
>
> H_2: The current alignment with the earth's magnetic pole of the magnetism of iron-bearing rock in any given region of the earth differs significantly from the alignment of the region's magnetic rock from earlier periods.

Laudan and Leplin further argue that the evidence which directly corroborates TC, namely, the data from remnant magnetism, *e*, supports H_2, but not H_1. However, since TC entails both H_1 and H_2, evidence directly corroborating H_2 indirectly supports H_1. In so arguing, however, they employ the two principles which Hempel showed cannot be jointly maintained without implying absurdities. Laudan and Leplin argue that *e* confirms H_2 and thus confirms TC. This is the so-called "converse consequence condition", which roughly says that evidence "flows up" the entailment relation. But Laudan and Leplin also maintain that *e* supports H_1, because TC entails H_1. This is Hempel's "special consequence condition". Now, Hempel does support this principle, but he argues that it cannot be universally used in conjunction with the converse consequence condition, since it would then follow that every statement supports every other one. The reasoning is reconstructed by Okasha (1997) as follows:

> [C]onsider any statement S. Every statement confirms itself, so S confirms S. By converse consequence, S confirms [S $\wedge$ T], since [S $\wedge$ T] $\Rightarrow$ S. By converse

> consequence, S confirms T, since [S $\wedge$ T] $\Rightarrow$ T. This result holds for arbitrary T. (p. 253)

What this shows is that the idea of indirect evidence cannot be argued for on the basis of formal entailment relations alone, which is exactly what Laudan and Leplin (1991) attempted.[116] Leplin, however, has more recently (1997a) attempted another argument against underdetermination, which evades Okasha's criticism. This time, he maintains that the thesis of underdetermination renders the notion of empirical equivalence empty. In other words, the empirical equivalence of two theories and their being underdetermined are incompatible claims. His argument runs roughly as follows: suppose that underdetermination is true and that there are two empirically equivalent rival theories, T_1 and T_2. Then, the empirical content of those theories is the same, that is, each theory in conjunction with all currently accepted auxiliary assumptions (A_1, for short) entails the same set of observation categoricals. But if underdetermination is true, then there should also be a rival set of auxiliary assumptions A_2, which is just as warranted by observations. Hence, the empirical content of any theory T remains indeterminate between (T and A_1) and (T and A_2), and the claim that two theories are empirically equivalent is unwarranted.

This argument does not have the problems raised by Okasha (1997), but, as Kukla has pointed out (1998, p. 65), it still fails for total theories, since the auxiliary assumptions are already part of the theory. Leplin (1997b) has responded to that criticism by claiming that the empirical equivalence of total theories can no more be established than the empirical equivalence of less than total theories. A total theory is a hybrid theory made up of an ordinary partial theory

---

[116] Okasha's criticism, it should be noted, does not apply to Massimi's analysis, because she does not conceive of indirect evidence as resting on purely formal entailment relations. Instead, she views it primarily in terms of a substantive interconnection between the hypothesis in question and auxiliary assumptions deemed warranted at the time.

plus the auxiliary assumptions warranted by the evidence available at the time. Leplin reasons that if underdetermination is true, "then there *are no* currently warranted theoretical auxiliaries from which to fashion the hybrid theories" (1997b, p. 213). For any one auxiliary assumption, there is a rival assumption which would be equally warranted by observations. Leplin also rejects the idea that we can just use whichever auxiliaries are believed at the time, regardless of whether they are in fact warranted. This, he claims, amounts to assuming that although underdetermination is true, it is not to be believed, which is an incoherent attitude (*ibid.*). This is true, but Leplin's argument is only effective against Kukla-like argumentative strategies. On the wider issue of underdetermination, it misses the point: the issue is not restricted to that of finding an appropriate set of auxiliary assumptions from which two empirically equivalent total theories can be constructed. Rather, it is whether we can rule out *a priori* the possibility of empirically equivalent total theories. Against that possibility, Leplin has offered no reasons. There surely are innumerable difficulties with the formulation of two such theories, as Hoefer and Rosenberg (1994) have pointed out, and also with the identification of those theories as empirically equivalent, as Quine has pointed out (see Chapter 4). Nevertheless, the possibility of its obtaining remains open.

Following a different line of reasoning than Leplin, Okasha (2002) has argued that the push towards total theories on the part of Kukla and others brings up another problem for those who want to defend underdetermination. The thesis of underdetermination, he argues, says that rival total theories can be equally warranted by the same observations and thus presupposes a distinction between theory and observations. Okasha, however, thinks that such a distinction cannot be drawn within total theories. He acknowledges that the distinction can be made relative to what a scientific community at a given moment takes to be observational and what it takes to

be theoretical. The borderlines would then shift according to the progress of science and technology, as some observations become increasingly theory-laden and technology enhances our perceptual capacities. But Okasha argues (2002, pp. 317-318) that the thesis of underdetermination, when applied to total theories, requires a principled distinction between theory and observation, one that is fixed and not relative to the state of science. Since no such distinction is available, the very notions of empirical equivalence and underdetermination become problematic.

Okasha's criticism, however, assumes that the fact that observations are theory-laden rules out the possibility of drawing the distinction between theories and observations within the theory itself. But it does not: from within a theory one can specify what counts as observations, and still admit that observations are insufficient to determine one theory above all others. It is difficult to imagine any theory/data distinction being drawn if not internally, from within a theory, especially if that theory is a total theory. Furthermore, the traditional qualms about where to draw the line between the observational and the theoretical, given the theory-ladenness of perception, need not concern us here. They can be altogether avoided if we switch from a distinction between theories and observations – which are very different sorts of things – to a distinction within kinds of sentences, as Quine proposes. Then the distinction is one among sentences only, observational and theoretical, and the thesis of underdetermination can be conceived as saying that all true observation sentences are insufficient to uniquely determine a theory. As we have seen in Chapter 2, Quine conceives a sentence as observational if it is an occasion sentence (true on some occasions and false on others), "on which the speakers of the language can agree outright on witnessing the occasion" (*PT*, p. 3). This acknowledgement of some sentences as observational does not imply that they are not theory-laden in any way.

Observation sentences are theory-laden insofar as the words they contain are also used in sentences of the theory which are not as tightly connected to observations. Yet, observation sentences remain observational insofar as the relevant community can immediately (that is, without further investigation) agree on its truth value upon witnessing the occasion described by the sentence.

5.4 **Conclusion**

We have just reviewed some of the most influential recent arguments for and against underdetermination. As we have seen, the debates have shifted toward total theories. This was due in part to the attacks on the claim that local theories can be empirically equivalent, and in part due to the claim that even if there are local theories that are empirically equivalent, those theories can be discriminated on the basis of how well they accord with more comprehensive theories that are also deemed warranted by the evidence available. This accords with Quine's views, for whom underdetermination is only interesting when formulated for "systems of the world". As a result of that shift, some of the arguments seeking to establish the empirical equivalence of local theories (for example, the arguments by van Fraassen and Stanford) have become secondary, since they at best can establish a thesis that is dependant on the underdetermination of total theories. In fact, the underdetermination of local theories by currently available evidence is just a corollary of the thesis of holism. Also, the shift towards total theories has rendered innocuous the criticisms of the notion of 'empirical equivalence' that are based on the claim that indirect evidence from auxiliary assumptions can *un*equally support local hypotheses. Even if that is true, it does not settle the issue of whether total theories can admit empirically equivalent rivals. Hence, some of the arguments for and against

underdetermination have become dependent on arguments which address primarily the thesis of underdetermination for total and global theories.

Nevertheless, the discussion has showed that the view of science as abiding by an oversimplified account of the hypothetico-deductive method naïvely and perhaps misleadingly contributes to the belief in underdetermination. But the thesis of underdetermination, if formulated as Quine and Hoefer and Rosenberg did, can survive even under a more intricate description of scientific method and practice. It is true that the thesis is then not much more than a conjecture, but it is nonetheless a conjecture which can hardly be denied, however unlikely its realization may be.

This review of the literature shows that the conjecture put forth by Quine and by Hoefer and Rosenberg remains as it always was: an open question. Stronger arguments for and against underdetermination have proven problematic at best. Kukla's purported proofs of the existence of empirically equivalent rivals for any theory require very loose criteria of theoryhood. Indeed, so loose that most practicing scientists would not even consider his alternatives. In fact, Kukla seems to accept any *logically* possible theory formulation as a legitimate theory. In so doing, he divorces the formulation of new theories from the analysis of the evidence, which is a move very much akin to those of a Cartesian skeptic. The end result is not distinguishable from radical skepticism, as some of Kukla's purported theories (most notably, T!) attest; hence, not an issue within philosophy of science properly speaking, but one about knowledge in general that deserves to be addressed elsewhere.

Criticism of the underdetermination of total theories have likewise failed. Leplin's charge that underdetermination is inconsistent with empirical equivalence presupposes that

empirical equivalence has to be established with the aid of auxiliary assumptions. This assumes we are not considering the underdetermination of total theories. Leplin, in fact, provides no reason to exclude the possibility of empirically equivalent total theories. Okasha's reasoning against underdetermination fails also: he assumes that the thesis requires an absolute theory/data distinction, but offers little reason why a more modest, Quinean distinction cannot work just as well.[117]

The conjecture of underdetermination implies that alternative descriptions of the world are in principle possible, and that observations may be ultimately insufficient to sort out which of them is true. It seems to remain an open question whether that conjecture also implies that one has to suspend judgment about the theoretical claims made by those theories (epistemic antirealism). We shall discuss this issue in the next chapter, in conjunction with Quine's assessment of the matter.

[117] To be fair, Okasha does devote a few lines to Quine. But he mostly defers the issue to a couple of pages by Friedman (1983, pp. 274-275) and a chapter by Churchland (1979, chapter 1), which would contain "telling objections to Quine's conception of observationality (Okasha, 2002, p. 318, n. 21). A more sympathetic reading of Quine on this matter might find in those mentioned pages a misrepresentation of Quine's views, rather than "telling objections". See, for example, Gibson (1988) and Hylton (forthcoming), as well as Quine himself: POS, *PT*.

6. REALISM AND UNDERDETERMINATION

6.1 Introduction

Philosophers who call themselves "realists" or "scientific realists" typically find problems in the thesis of underdetermination. Examples can be found in the works of Laudan, Leplin, Kitcher, and Norton, mentioned in the previous chapter. Philosophers who believe in the thesis of underdetermination, on the other hand, usually deny one aspect or another of scientific realism. The works of van Fraassen, Bergström, Kukla, and Stanford – also mentioned in the previous chapter – illustrate the point. On this issue, Quine's views do not line up with either camp: he asserts his realism and his belief in the thesis of underdetermination. This aspect of Quine's philosophy has been particularly difficult for his readers to understand, and it has not only incited objections from commentators, but also left Quine himself pondering on what to say about the truth of empirically equivalent rival theories, as we saw in Chapter 3. Quine apparently remained undecided between an ecumenical and a sectarian attitude in regards to the truth of rival theories, and only hesitantly favored the sectarian attitude in his last published writing on the matter, *PT* (pp. 95-102). The reasons that pulled him in each direction have important similarities with the motivations which drive philosophers toward realist and antirealist accounts of science, as we shall see.

Neither the sectarian nor the ecumenical attitudes entertained by Quine, however, fall on the antirealist side of the debate, since he never backed down on the belief that we cannot but regard the theories we happen to use as true. The question for him was not whether *any* of the empirically equivalent rival theories can be regarded as true, but whether *more than one* theory can be considered true, and whether that makes much difference in the case of empirically

equivalent theories. Nonetheless, the main motivation for the ecumenical attitude in Quine's work is empiricism. But in Quine's work the question about whether theories that rival our own can be regarded as true takes place within the context of an unquestioned naturalism. Even empiricism, for Quine, is based on the naturalist view of science. So in this sense, Quine's motivation for the ecumenical attitude, empiricism, cannot have the same consequences that it has in other philosophers. On the other hand, Quine's main motivation for the sectarian attitude is naturalism, and the associated realism about the objects posited by our best theories. His justification for the sectarian attitude is the belief that we cannot but assert the truth of the theories we use – even while acknowledging its fallibility – and that while using a theory we cannot but reject the truth or meaningfulness of its rivals, even if we acknowledge the alternative points of view that engender them as legitimate and perhaps fruitful.

Most authors are not sympathetic at all to Quine's philosophical torments on the matter, however. Even among some of his most charitable readers, such as Gibson (1986) and Bergström (1984 and 1993), the charge has been made that underdetermination brings out an inconsistency within his philosophy, or that there is something unintelligible about it. Others simply disregard Quine's self-description as a "robust realist" (TTPT, p. 21), and suggest that he is in fact an antirealist.[118] Our goal in this chapter is to understand what is it about Quine's philosophy that allows him to reject the arguments from underdetermination to antirealism which we reviewed in the previous chapter, and to hold on to realism and underdetermination at the same time, even if hesitating about what to say of the truth of empirically equivalent rival theories. We have already seen in Chapters 2 and 3 that Quine's thesis of underdetermination is substantially weaker than most other versions of the thesis currently debated in the literature.

---

[118] For example, Sarkar (2000, p. 189).

This weakened thesis may perhaps explain why underdetermination was never thought by Quine to pose any serious challenge to realism. But realism as Quine understands it could survive even a stronger thesis of underdetermination, as we shall see below. Quine's conception of realism itself is also relatively weak in the sense that it does not have most of the metaphysical implications traditionally associated with the notion. Section 6.2 below briefly describes Quine's realism; section 6.3 elaborates a Quinean response to arguments for epistemic antirealism and to arguments against underdetermination.

6.2 **Quine's realism**

It is not hard to find passages in Quine's works where he describes himself as a realist.[119] But those are rather brief passages, and realism is not the primary concern of the texts where they figure. There is indeed a sense in which this was never an issue for Quine. As a metaphysical thesis asserting the existence not merely of the unobservable entities posited by science, but of physical entities in general, Quine takes realism to be not just true, but the only intelligible position we can entertain. We can and certainly do question the nature of the physical entities posited by the science of any given historical period, and even what to count as *physical*, or as an *entity*. These are some of the questions that drive the scientific enterprise broadly conceived. Quite another matter is the skeptical question about the existence things outside our minds, and of physical things in general:

> We cannot significantly question the reality of the external world, or deny that there is evidence of external objects in the testimony of our senses; for, to do so is simply to dissociate the terms 'reality' and 'evidence' from the very applications which originally did most to invest those terms with whatever intelligibility they may have for us. (SLS, p. 229)

---

[119] For example, SLS, pp. 229-230; PR, pp. 250-252; *WO*, §§ 1, 6, and 48; and TTPT, pp. 21-22.

The questions about which objects there are, and whether we can infer the existence of unobservable entities from what we do observe, are certainly part of the pursuit of truth in the sciences. But to question the existence of objects in general, "the reality of the external world", is for Quine something of an illusion, an attempt to make sense of words while forgetting what they mean. The very notion of 'reality', Quine argues (WO, p. 22), acquires its meaning and utility in the context of the evaluation of evidence for a given claim. The motivation only makes sense under the assumption that there are some undeniable instances of things we consider real and existent. In the absence of those, the notion of reality could never have played a useful role. It would not, in other words, have found its way into our language: "Unless we change meanings in midstream, the familiar bodies around us are as real as can be; and it smacks of a contradiction in terms to conclude otherwise." (PR, p. 251)

In the context of evaluating whether, say, a suspicious-looking dollar bill is forged, one might naturally ask whether it is *real*. Similarly, when witnessing a very unusual or unfamiliar event, say, the aurora borealis, one might also ask whether what is seen is real. But asking whether what we see is real in these contexts relies on some criteria for sorting out the real from the unreal (or imagined or fake or artificial). A specialist from the U.S. treasury, for example, will certainly be able to sort out legal currency from forged dollar bills. Likewise, the trained eyes of an astronomer will be capable of telling apart the aurora borealis from a show of artificial lights put up by a movie studio. 'Real' in these contexts is contrasted with 'fake', 'forged', 'artificial', and the like, and what sets them apart are observable marks: "the testimony of our

senses does (contrary to Berkeley's notion) count as evidence (…), such being (…) just the sort of thing that evidence is" (*ibid*.).[120]

But there is also, at least among philosophers, a sense of 'real' which it is not connected at all to empirical criteria. The skeptical question about the reality of the external world is an example. It enjoins us to entertain the possibility that even those things which we take to be paradigmatically real may in fact be unreal. For Quine, this is not an intelligible conjecture, precisely because it employs the notion of 'reality' and related notions in a way that deprives "them of the very denotations to which they mainly owe such sense as they make to us" (*WO*, p. 22).[121] It dissociates 'reality' from the empirical contexts on which it first acquires its usefulness.

Quine does not separate into different categories the ordinary, common sense notion of 'reality' from its perhaps more technical use in the sciences. Science is in general a continuation of common sense, albeit under more systematic and self-conscious constraints. The general purpose and scope of science is narrower, however:

> The general task which science sets itself is that of specifying how reality "really" is: the task of delineating the structure of reality as distinct from the structure of one or another traditional language (except, of course, when the science happens to be grammar itself). The notion of reality independent of language is carried over by the scientist from his earliest impressions, but the facile reification of linguistic features is avoided or minimized. (SLS, p. 233)

The idea that reality is independent from theory is for Quine something our own theory of the world affirms, not some truth of a higher or transcendent sort. In fact, for Quine there are no such higher truths. Quine's naturalism is thoroughgoing, he describes it as "the recognition that

---

[120] See also *WO*, p. 3.

[121] Similar views on this matter can be found in Austin (1962) and Cavell (1979).

it is within science itself, and not in some prior philosophy, that reality is to be identified and described" (TTPT, p. 21). Naturalism is the acknowledgment that we cannot speak about the world but from within some scientific theory or other. There is no neutral standpoint outside all theories from which reality can be conceived and the truth of a theory ascertained. We ourselves are part of reality, as are our attempts to accurately describe ourselves and the world around us. The very notion of 'reality' is a notion within a theory, in fact a very basic notion within our own general theory of the world.

Compared to so-called "metaphysical realism", Quine's realism may sound very modest indeed. Metaphysical realism assumes the possibility of an absolute epistemic gap between the theories and what theories purport to describe: one that cannot be bridged even by a scientific theory that is fully adequate. A metaphysical realist might be willing to give more credit to radical skepticism for that reason. The idea that no matter how well theories conform to observations, they might still be false (the evil genius hypothesis of Descartes, for example), is something Quine dismisses outright:

> Our scientific theory can indeed go wrong, and precisely in the familiar way: through failure of predicted observation. But what if, happily and unbeknownst, we have achieved a theory that is conformable to every possible observation, past and future? In what sense could the world then be said to deviate from what the theory claims? Clearly in none, even if we can somehow make sense of the phrase 'every possible observation'. Our overall scientific theory demands of the world only that it be so structured as to assure the sequences of stimulation that our theory gives us to expect. More concrete demands are empty, what with the freedom of proxy functions. (TTPT, p. 22)

This view of science is an aspect of what Quine calls "epistemology naturalized", which conceives the traditional epistemological questions about the justification of knowledge in general as an outgrowth of more ordinary scientific inquiry. The challenge for Quine is that of providing a naturalistic account of knowledge in general, and of science in particular – "if our

science were true, how could we know it?" (*RR*, p. 2) – that does not rely on any sort of absolute or external grounding of the sciences in a more fundamental kind of knowledge. As he puts it in *WO*, "we can never do better than occupy the standpoint of some theory or other, the best we can muster at the time" (p. 22). Even the most fundamental epistemological questions thus must be tackled from within our general theory of the world.

The traditional metaphysical questions about the nature of reality are thus for Quine tasks of science itself. In its broadest features, he argues, the "limning of the most general traits of reality" (*WO*, p. 161) is not to be distinguished from the task finding the "simplest, clearest overall pattern of canonical notation" (*ibid.*), that is, the basic elements of the language of science itself. Metaphysics, for Quine, can thus be seen as evolving hand in hand with epistemology. While describing the epistemological process of theory building, he wrote, the objects posited by science are just that: posits. But from the point of view of the theory being built, those posits are our ontology, insofar as the theory is sufficiently warranted: "To call a posit a posit is not to patronize it." (*WO*, p. 22)[122]

Science, here, is conceived very broadly. It is not merely another name for physics, nor does it presuppose physicalism, the view that only the objects posited and required by physics exist. On Quine's view, science is not essentially tied to physics or to physicalism, although he did argue that physicalism is what a regimented theory of the world now commits us to. But there is nothing intrinsic to the scientific practice which should prevent it from abandoning physicalism, if the evidence pointed in that direction:

> The science game is not committed to the physical, whatever that means. Bodies have long since diffused into swarms of particles, and the Bose-Einstein statistic

[122] For further discussion of Quine's realism, see Gibson (1988) and Hylton (1994, 1997, and 2004).

> (…) has challenged the particularity of the particle. Even telepathy and clairvoyance are scientific options, however moribund. It would take some extraordinary evidence to enliven them, but, if that were to happen, then empiricism itself – the crowning norm (…) of naturalized epistemology – would go by the board. For remember that that norm, and naturalized epistemology itself, are integral to science, and science is fallible and corrigible. (*PT*, p. 21)[123]

Quine's general naturalistic account of science acknowledges that even our best theories of the world can be wrong. But it denies that we can assess the truth of a theory from outside all theories. This in part explains why his vacillations between the ecumenical and the sectarian attitudes discussed in Chapter 3 were never thought (by Quine himself, that is) to cast doubt on the idea that to some degree we can and do identify and describe reality.

6.3 **Quine's responses to concerns about underdetermination**

As we have seen in the previous chapter, the thesis of underdetermination has recently been challenged for two main reasons. The first is that it is thought to be confused or incoherent by epistemic realists. The second is that underdetermination is thought to conflict with the ever more accurate predictions of natural science seem to allow. We shall deal with each in turn.

As we have seen in Chapter 3, Quine's thesis of underdetermination, as it came to be formulated by Quine from 1975 onwards, states that empirically equivalent rival systems of the world are possible, such that if we were to discover them, we would systematically fail to render them logically equivalent through translation. The concern that this raises for the epistemic realist is that in the event we were to find two such systems, we would be unable to assert the

---

[123] Quine's view here seems to be that empiricism would "go by the board" because it is conceived as only allowing for our five senses as sources of information about the world. Other philosophers – for example, William James and C.D. Broad – have questioned these restrictions, and argued that what we today might call "extra-sensory perception" is still a kind of perception, and that allowing for it does not have to be seen as contradicting empiricism.

truth of either. Bergström (1993), for example, argues that the only rational attitude in such cases is agnosticism.[124]

Suppose that we find two theories which we somehow manage to recognize, even if tentatively, as empirically equivalent. Suppose, additionally that all attempts at intertranslation fail. The two theories are thus thought to be equally warranted by observations but to present alternative descriptions of the unobservable world. The two theories might even imply incompatible sentences, say, $(\forall x)Fx$ and $\sim(\forall x)Fx$, respectively, as we saw in Chapter 3. But, following the "trivial expedient" proposed by Davidson (1990) and adopted by Quine (TI, *PT*), one of the theories can be replaced by another which does not contain the troublesome predicate. The new theory could, for example, replace it with a predicate which is new to both rivals. Suppose that F', is the new predicate. Then the original inconsistency will disappear, since the two theories will now imply $(\forall x)Fx$ and $\sim(\forall x)F'x$, respectively. By systematically repeating this same procedure, eventually all incompatibilities can be avoided. To be sure, the original theories are still incompatible. But since it can only occur at the level of theoretical predicates whose extension cannot be fixed ostensively, there is no reason not to treat the incompatibility of the original theories as resting on an equivocation, and use the "trival expedient" to work around it.

The remaining conflict that the new pair of theories can be thought to have is one of non-intertranslatability. The theories may contain terms and sentences, perhaps even engendered by the trivial expedient, such that the sentences of which they are part cannot be rendered logically equivalent through translation. Now suppose that one of those theories implies

(1) $(\forall x)(F_1x \vee F_2x \vee \ldots \vee F_nx)$,

[124] See also van Fraassen (1980).

and that its rival implies

(2) $(\exists x)Gx$.

Then, although the two theories do not imply logically inconsistent sentences, there still seems to be some sort of incompatibility, which we could perhaps call "semantic". Each of the theories seems to posit a different set of entities. One theory says that everything is an *F* of one kind or another, whereas its rival says that there is at least one *G*. Surely, the argument goes, there is an inconsistency here. The two theories cannot both be true: If one theory says that everything is an *F* of one kind or another, then there are no *G*s. But the rival theory says precisely that there is at least one *G*. So they cannot both be true. And since we cannot know which of the two is true, given their empirical equivalence, we should suspend judgment altogether about the theoretical entities and principles posited by both theories.

This has been put forth as a major reason for agnosticism – for example, by Bergström (1984, 1993). The reasoning, however, relies on a problematic inference. The argument assumes that the theory which affirms that $(\forall x)(F_1x \vee F_2x \vee \ldots \vee F_nx)$ also implies that there are no *G*s while acknowledging that the predicate *G* does not belong to the vocabulary of that theory nor can be defined in terms of the vocabulary of that theory. In other words, the argument assumes that $(\forall x)((F_1x \vee F_2x \vee \ldots \vee F_nx) \rightarrow \sim Gx)$, but also acknowledges that this last sentence cannot belong to any theory, since the predicates which comprise it belong to rival theories which have mutually exclusive and non-intertranslatable vocabularies. But from the perspective of the latter theory, *G* is meaningless and useless. Admitting such a term into the theory would add nothing to the empirical content of the theory, and would detract from its simplicity and elegance:

> It is as if some scientifically undigested terms of metaphysics or religion, say 'essence' or 'grace' or 'Nirvana', were admitted into science along with all their pertinent doctrine, and tolerated on the ground merely that they contravened no observations. It would be an abandonment of the scientist's quest for economy and the empiricist's standard of meaningfulness. (RRG, p. 157)

Bergström (1993) also argues that the ecumenical attitude entertained by Quine at various times is unintelligible, because if theories are rivals then it is incoherent to acknowledge both as true. It is hard not to give some credit to Bergström on this point, and it seems that Quine eventually rejected the ecumenical attitude and settled instead – although rather hesitantly – for the sectarian attitude, for reasons similar to the ones listed by Bersgström. But Bersgström contends that also sectarianism is unjustified: given the equal warrantedness of the rival theories we cannot assert the truth of either. Only agnosticism would be justified.

This is a point on which Quine asserts his naturalistic view of science. He allows no room for doubt about a theory which thoroughly conforms to all observations. If we manage to find a global theory which is empirically adequate, then there is no justification for doubt. Doubt about the truth of a theory can only come from the failure of some predicted observation. In the absence of those, either we adopt what seems to be the best theory available and say that it is true, or we adopt an empirically equivalent alternative and say that the alternative is true. Quine did vacillate, as we have seen, on whether we should say that all empirically equivalent rivals to our theory are true also. Here it is unclear what to say. In any case, the justification for the sectarian attitude is not that we can be certain of the truth of the theories we happen to hold at any given time. The theories we hold may be false; our science is indeed fallible. Instead, the justification for sectarianism is that we cannot but adopt one theory or another, and from the perspective of that theory alternatives will be false or partially meaningless. Furthermore, at least some fundamental aspects of the theories we adopt cannot be questioned while we use those

theories. To be sure, which aspects of a theory we hold fast to may vary in time. But for some aspects of a general theory to be questioned, others must remain unquestioned. And about the latter aspects of a theory, belief and use go hand in hand.[125] At an extreme case, to use a theory to make predictions of observations *is* to believe, even if tentatively, in what the theory implies. This, however, is hardly a reason to question epistemic realism at all. The fact that our beliefs are fallible does not by itself entail that one cannot have good reasons for asserting the truth of whichever theory seems best warranted by the evidence available.

As we saw in Chapter 5, a second major concern about underdetermination in the recent literature has been that it relies on an overly simplified account of scientific method. This seems right, but it need not be seen as a sufficient reason for rejecting the thesis of underdetermination as it was formulated by Quine. Quine's thesis is compatible with the inductive methods of reasoning used by scientists to develop new theories, and the certainty with which most scientific theories are asserted by members of the scientific community is neither supposed to entail that the results of science obtained so far are infallible, nor that there cannot be rival theories which account for the same observations equally well. In particular, even under a more careful description of scientific practice, the conjecture of empirically equivalent global rivals that are not intertranslatable remains open, despite its unlikelyhood:

> Might another culture, another species, take a radically different line of scientific development, guided by norms that differ sharply from ours but that are justified by their scientific findings as ours are by ours? And might these people predict as successfully and thrive as well as we? Yes, I think that we must admit this as a possibility in principle; that we must admit it even from the point of view of our

[125] For discussion of related issues, see Horwich (1991), and Hacking (1984).

> own science, which is the only point of view I can offer. I should be surprised to see this possibility realized, but I cannot picture a disproof. (*TT*, p. 181)

This, however, is only a possibility in principle. One that may perhaps never be realized. Furthermore, it is a possibility that seems so far removed from our current practices in theory building that it should at least remain an open question whether, if such an event were to happen, we would not want to reconsider not just the truth of the theory we happen to hold, but also the very methods of assessing the truth of theories, perhaps even the rationale for sectariansm:

> The fantasy of irresolubly rival systems of the world is a thought experiment out beyond where linguistic usage has been crystallized by use. No wonder the cosmic question whether to call two such world systems true should simmer down, bathetically, to a question of words. Hence also, meanwhile, my vacillation. (PT, p. 100-101)

7. CONCLUSION

This dissertation contains an analysis and a defense of the plausibility of Quine's thesis of underdetermination. It was shown in the preceding chapters that his views have been frequently misread, and that the thesis of underdetermination he put forth is substantially weaker than most of what goes by that name in the recent literature. In particular, Quine's thesis does not assert the existence of empirically equivalent rival "systems of the world". Quine views the existence of such systems, or theories, as a mere conjecture. He argues that not only is it logically possible that rival global theories may exist that predict exactly the same observation categoricals, but also that holism and the "less-than-rigid connections that obtain between sensory stimulus and physical doctrine" (PR, p. 254) suggest that that possibility could obtain, however unlike it may seem given the current state of science. But holism alone is insufficient to establish the underdetermination of global theories. For Quine it is an "open question" (EESW, p. 327) whether a global theory can be underdetermined and admit rivals that are equally warranted by observations. Perhaps global theories have constraints that are yet to be discovered which rule out alternative accounts. Perhaps all empirically equivalent global theories are in fact intertranslatable versions of the same theory. Perhaps there are no empirically adequate global theories to begin with. All of these possibilities are in principle admitted by Quine. None of them is be ruled out by what we currently know about reality and about the construction of theories in the sciences.

Quine's thesis of underdetermination makes instead a milder claim. It says merely that there are empirically equivalent global theories which, "if we were to discover them, we would see no way of reconciling by reconstrual of predicates" (EESW, p. 327), or translation (NNK, p. 80). This thesis, however, does not in principle rule out the possibility that all pairs of

empirically equivalent global theories that there are (in the platonic sense, so to speak) are in fact intertranslatable versions of the same theory. That is, it does not rule out the possibility that for any two empirically adequate and empirically equivalent global theories, there is (also, in a platonic sense) a manual of translation which renders the two theories logically equivalent. Quine's thesis of underdetermination does not deny that that is possible. It says instead that there are (again, in a platonic sense) empirically equivalent global theories which, if we were to discover them, *we would systematically fail to find a manual of intertranslation*. This could be due to the fact that the two theories are so radically different that any manual of intertranslation would be exceedingly complex to be mastered by human beings. But it could also be due to the fact that there are no manuals of intertranslation at all for some of those theories. Which one of these two alternatives is the case, is something left open by Quine's thesis.

These important subtleties of Quine's version of the thesis have been largely ignored in the literature on underdetermination. In that literature, the thesis has been either understood as making a stronger claim to the effect that there are equally warranted rivals to any given theory, or the weaker claim (also known as "transient underdetermination") that local theories are underdetermined by currently available evidence. Both of these versions of the thesis have been taken to be incompatible with scientific realism in one way or another. Hence, philosophers of science who are inclined towards realism have systematically criticized the thesis of underdetermination, and frequently extended those criticisms to Quine himself. On the other hand, philosophers who are inclined to scientific antirealism have often relied on the thesis of underdetermination to justify their disbelief or doubt in the theoretical claims of contemporary science.

The present dissertation attempts to clarify these issues by putting forth an analysis both of Quine's views and of the various versions of the thesis of underdetermination as it is understood in the current literature. The results achieved can be summarized as follows:

1. The thesis of underdetermination proposed by Quine is suggested by holism – the doctrine that the empirical content of a theory cannot be sorted out distributively among its individual sentences – but not entailed by it. Holism does entail that alternative modifications to a theory can engender new theories which do not imply an observation categorical that is known to be false. But the revised theories might not be empirically equivalent, and typically they are not. Hence, the revised theories might be unequally warranted by observations currently available or available in the future. Furthermore, even if all the revised theories were empirically equivalent, holism alone does not entail that they will be sufficiently different to count as rival theories. It could so happen that all the revised theories end up being intertranslatable versions of the same theory, that is, that the theories could be rendered logically equivalent by some (known or unknown) manual of translation, in which case the alternative modifications entailed by holism would be rivals only in a trivial sense.

As we have seen in Chapter 2, this is a point on which Quine is oftentimes misread. He is frequently taken to have asserted that the thesis of underdetermination is grounded on holism. Sometimes Quine's thesis of holism and underdetermination are confused, or taken to say essentially the same thing. This is a gross misunderstanding of Quine's work. Holism is a fundamental thesis in his philosophy in the sense that it has implications for almost all other Quinean doctrines – for example, the theses of indeterminacy of translation and ontological relativity, Quine's criticism of Carnap's use of the analytic/synthetic distinction, Quine's epistemology naturalized and his view of philosophy as continuous with science and of science

as continuous with common sense. Holism is also fundamental to Quine's philosophy in the sense that he took it to be obviously and trivially true. Underdetermination, on the contrary, does not carry with it significant implications for most of Quine's doctrines, and it is certainly not obviously or trivially true. In fact, as we have seen, the thesis itself is somewhat of a conjecture.

2. One of the reasons why Quine's thesis of underdetermination has been misunderstood is the fact that it is weaker than other versions of the thesis currently debated. Another reason is the fact that Quine changed his own formulation of the thesis twice, and vacillated about its implications for the truth of rival theories. Chapter 3 has tracked down those changes and vacillations in Quine's views and argued that they can be mapped onto two distinct sets of texts. Quine's writings on underdetermination up until 1975, when EESW was published, are primarily concerned with the formulation and justification of the thesis. The texts published thereafter, on the other hand, mostly presuppose the formulation put forth in EESW and discuss whether, given that thesis, we should regard all theories that are empirically equivalent to our own as true, or only our own theory as true.

The two sets of issues are distinct but interconnected. The changes that Quine's formulation of the thesis underwent were two: His initial formulation (PR, *WO*) speaks rather vaguely of *alternative* theories that are equally warranted by observations. Afterwards (RC, RIT) Quine speaks of empirically equivalent theories that are *logically incompatible*. His final formulation of the thesis (EESW, NNK, RC, *PT*) drops the notion of 'logical incompatibility' and speaks instead of theories that cannot be rendered logically equivalent by reconstrual of predicates or by translation. All of Quine's writings published after 1975 use the latter formulation of the thesis. These later texts no longer address the issue of how to formulate and justify the thesis, but a more specific issue: that of the truth of theories that rival our own. This

only became an issue for Quine once he dropped the notion of 'logical incompatibility' from his formulation of the thesis. Clearly if underdetermination implies that there are empirically equivalent theories that are logically incompatible, then, by logic alone, at most one can be true. By definition, sets of sentences are incompatible if they cannot all be true at the same time. Once the notion of 'logical incompatibility' was dropped from the formulation of the thesis, however, the possibility was open for empirically equivalent rival theories to be simultaneously regarded as true. This was the primary issue addressed by Quine in his writing on underdetermination after 1975.

If no distinction is made between these two sets of issues (and the corresponding texts), then it may seem that Quine's thesis of underdetermination has some substantial metaphysical or epistemological consequences. In particular, it may seem that Quine's vacillation about whether empirically equivalent rival theories can be simultaneously true are similar in nature to the debates on whether underdetermination is a good reason for antirealism. This may have contributed to the somewhat diffused understanding of Quine's thesis of underdetermination as having metaphysical and epistemological implications similar to those debated in the current philosophy of science literature on underdetermination.

3. An objection to Quine's thesis of underdetermination is that of incoherence. If two theories are empirically equivalent, the objection goes, then in principle they should be intertranslatable. This seems to be supported by Quine's thesis of indeterminacy of translation, which has indeed led some commentators to think that Quine's theses of indeterminacy and

underdetermination are incompatible.[126] Quine does sometimes say – for example, in RIT (pp. 179 ff) – that indeterminacy of translation follows from underdetermination, since the only constraint on translation is agreement on observation sentences. A manual of translation is underdetermined by the observable speech behavior of the native speakers, hence alternative manuals exist that translate the foreign speech into different native idioms. But then it would seem that any two empirically equivalent theories should be intertranslatable, since they must agree on all observations. This is perhaps the kind of reasoning – suggested by Quine himself at times – which seems to motivate the confusion about indeterminacy and underdetermination. The two theses make very different claims, however. The thesis of underdetermination says that there are theories which are empirically equivalent and for which we may systematically fail to render logically equivalent through translation. The thesis of indeterminacy of translation, on the other hand, says that there are alternative manuals of translation compatible with any given set of observable behavior of the native speakers of a language. The lack of intertranslation which the thesis of underdetermination alludes to is a lack of intertranslation of the theoretical, or non-observational, terms and sentences of a global theory, whereas the manuals of translation alluded to by the thesis of indeterminacy provide alternative translations of the *observable* behavior of the speakers. The latter thesis remains true even if theories were *determined* by observations. Even if we could be sure that only one physical theory is allowed by observations, we could still translate what the speakers of a foreign language say in divergent manners. What they mean by what they say is left indeterminate, even if physics is not:

> The indeterminacy of translation differs from the underdetermination of science in that there is only the natives' verbal behavior for the manuals of translation to be right or wrong about; no claims are made regarding hidden neural mechanisms. If

[126] See Bechtel (1980), Dummett (1981), Gibson (1986), Gemes (1991), and Yalçin (2001); see also Quine's ITA.

> translators disagree on the translation of a Jungle sentence no behavior on the part of the Jungle people could bear on the disagreement, then there is simply no fact of the matter. In the case of natural science, on the other hand, there is a fact of the matter, even if all possible observations are insufficient to reveal it uniquely. The facts of nature outrun our theories as well as all possible observations, whereas traditional semantics outruns the facts of language. (ITA, pp. 9-10)

The difficulty that the theses of indeterminacy and underdetermination have brought to Quine's readers is there are some formulations of the thesis of indeterminacy which suggest that the only constraint on translation is agreement on observation sentences. But that seems then to entail that any two empirically equivalent theories would be intertranslatable. Although there are passages in Quine's work in which he does characterize indeterminacy of translation in that way, his final or considered view on the matter was that there are also other constraints on translation, in particular the *structure* of the theories translated. A translation which translates the whole of a theory but not its parts would be rather queer indeed (see EN, pp. 79-80).[127]

These were issues most notably discussed by Gibson (1988, 1996, and 1998), and only briefly touched upon in Chapter 4. But Gibson's discussion left open the question of what Quine could have meant by lack of intertranslatability between theories. It might seem from some of Quine's writings on indeterminacy of translation that any two empirically equivalent theories are *ipso facto* intertranslatable, and that the conjecture which the thesis of underdetermination invites us to entertain might in fact be empty, or impossible. Chapter 4 contains an argument in support of the intelligibility of Quine's thesis. It describes underdetermination as a limiting case of Kuhn's thesis of incommensurability, which says that rival theories in the history of science have been put forth that contain terms that cannot be incorporated into the language of one another.

---

[127] Incidentally, this might be one of the reasons why Quine came to speak in his papers in the 1980s and 1990s of holophrastic indeterminacy (or indeterminacy of sentences) as a conjecture, as opposed to the indeterminacy of reference (or ontological relativity), for which he had a proof (see RJW, 728).

Notorious examples are those of "phlogiston", which has no counterpart in contemporary chemistry, and "neutrino", which has no counterpart in eighteenth century physics.

It has been objected (originally by Putnam and Davidson) that the idea of untranslatable terms and sentences is incoherent. If we manage to identify something as a term or a sentence then presumably we can describe what role it plays in a language, and thereby translate it. This is not true, however: to describe the role a term plays in a language is not to translate it. Some terms and sentences are such that they remain foreign even after a description of their original use was given. Such is the case "phlogiston" mentioned above. We can describe its use, but not actually incorporate it into the language of the theories. There is a sense in which we might want to acknowledge that description of the use of the term as a "translation". But there is also a sense in which that is no translation at all: a description gives us the rules of its use in the foreign language, not the translation itself. The distinction to be made here is one familiar to bilinguals: to be able to speak a foreign language does not entail being able to translate each of the spoken sentences and terms of that foreign tongue into sentences and terms of one's native tongue. What historians of science typically do is to teach us enough about the older theories so that we become sufficiently fluent in the use of its expressions to make sense of what they meant. Understanding them, however, does not entail the capacity to translate. These were points originally made by Kuhn and Feyerabend in response to the objections by Putnam and Davidson to the thesis of incommensurability. Chapter 4 showed that those points can also be adapted to work as responses to similar objections to the thesis of underdetermination.

The idea that the empirically equivalent rival theories are in principle intertranslatable rules out by definition the possibility that those theories will not contain terms and phrases which are just as impermeable to intertranslation as some of the examples discussed by Kuhn and

Feyerabend. These examples, of course, are not proofs of the thesis of underdetermination, since they are taken from theories that are neither empirically equivalent nor empirically adequate. However, they do show that cases of untranslatability are in principle possible. And by doing so those examples shed some light on the kind of rivalry which underdetermined theories can be thought to have, and thus lend some additional credence to Quine's thesis of underdetermination.

4. In the recent literature the thesis of underdetermination has been formulated in various ways. Philosophers who defend the thesis tend to adhere to weak formulations, while those who reject it usually formulate it in stronger terms. The general result is that the recent debates on the thesis have suffered greatly from equivocation.

A very weak thesis has been proposed by Sklar (1981) and Stanford (2001), sometimes dubbed "transient underdetermination". It says that local theories are underdetermined by currently available evidence. The evidence available for the theories we held in the past support just as well most of the theories that succeeded them. The evidence available for Newtonian physics during eighteenth century, for example, also supports contemporary physics. In general, there is no reason not to expect that the theories we currently hold are not likewise underdetermined by the evidence currently available. New theories yet to be discovered are surely at least as warranted by the body of evidence now available. This is a very weak version of the thesis of underdetermination. So weak, in fact, that it is not clearly distinguishable from the thesis of holism. Most, if not all, the theories we now hold are the result of modifications made to prior theories in light of counterexamples. But holism says that various alternative modifications can prevent a theory from implying an observation categorical that is known to be false. It is true that the result of those modifications are mostly not empirically equivalent

theories. But then again, the theories which illustrate the thesis of "transient underdetermination" are not empirically equivalent either.

Another formulation of the thesis is that of Laudan and Leplin (1991), which says that theories in general admit empirically equivalent rivals that are equally warranted by the evidence. This is a much stronger formulation of the thesis; so strong, in fact, that it can hardly be satisfied at all. Laudan and Leplin purport to have "defeated" it by showing that it cannot hold in general: local theories, in particular, can be empirically equivalent in the sense that they entail the same observation categoricals and yet be unevenly warranted by indirect evidence. The argument, however, only holds for theories which are less than total, and hence admit indirect evidence from auxiliary assumptions.

Yet another formulation of the thesis of underdetermination says that any theory has empirically equivalent rivals. This seems to be a stronger formulation of the thesis, but its strength depends on what one counts as a theory, and on how one defines 'empirical equivalence'. Kukla (1998) argues that his version of the thesis can be proved, even for global theories. For any given theory one could devise an appropriate algorithm for constructing a new theory which is empirically equivalent to the original. One example he offers is that of a rival theory which contains all the observation categoricals implied by the theory out of which it was created and the negation of the conjunction of all the other sentences. These algorithms suggested by Kukla, however, reveal that he adopts a very lax standard on both what to count as a theory and what to count as a *rival* theory. He does not even seem concerned to rule out theories which can be rendered logically equivalent by translation, such as theories TN(0) and TN(v) mentioned in Chapter 5 reveal. The thesis of underdetermination he thus proposes admits as relevant for science some alternatives which only a radical skeptic would consider. By

weakening some of the requirements of the thesis, Kukla is able to "prove" it, but only at the cost of making it irrelevant for the concerns of scientists and philosophers of science.

These three examples illustrate how some of the most influential discussions of the thesis in the recent literature use formulations of the thesis which turn out to be undesirable for some reason or another. They also show how insightful Quine's discussion of the thesis was. The reasons offered against underdetermination in the recent literature leave mostly unscathed Quine's "vague and modest" version of the thesis; the reasons purporting to prove or demonstrate the thesis tend either to leave it as weak as holism, or to relax the requirements it needs to satisfy, and thereby leave it nearly indistinguishable from radical skepticism.

5. Quine's formulation of the thesis of underdetermination left a problem open: what to say about the truth of empirically equivalent rival theories. As we have seen in chapters 3 and 6, this was an issue about which Quine vacillated between an ecumenical and a sectarian attitude. Unlike recent accounts of underdetermination, neither of these was thought by Quine to be incompatible with realism. Some objections have been put forth against Quine's view. One of them says that the ecumenical attitude is unintelligible, because it enjoins us to admit as true theories which we also think present conflicting accounts of reality, and that sectarianism is unjustified given the equal warrantedness of rival theories. Bergström (1993), who argues along these lines, claims that the only rational attitude is agnosticism. Given the equal warrantedness of rival theories, we must suspend judgment as to their truth, even if they agree on all observations. Alternative theories cannot be both true, even if they only diverge on the theoretical entities and principles they posit. But the argument for agnosticism requires that belief in a theory be separable from the adoption of that theory and its use for making predictions

of observations. This is possible to some extent, but only while some aspects of a theory are not put into doubt. As a strategy to be pursued in general, it does not seem feasible.

REFERENCES

Ariew, R.: The Duhem thesis. British Journal for the Philosophy of Science, 35: 313-325, 1984.

Austin, D.: What's the Meaning of "This"? Ithaca, NY: Cornell University Press, 1990.

Austin, J.L.: Sense and Sensibilia. Oxford: Oxford University Press, 1962.

Bain, J.: Weinberg on QFT: Demonstrative induction and underdetermination. Synthese, 117: 1-30, 1999.

Bar-On, D.: Semantic indeterminacy and scientific underdetermination. Pacific Philosophical Quarterly, 67: 254-263, 1986.

Bechtel, P.W.: Indeterminacy and underdetermination: are Quine's two theses consistent? Philosophical Studies, 38: 309-320, 1980.

Ben-Menahem, Y.: Equivalent descriptions. British Journal for the Philosophy of Science, 41: 261-279, 1990.

Benacerraf, P.: Mathematical truth. Journal of Philosophy, 70: 661-78, 1973.

Bergström, L.: Underdetermination and realism. Erkenntnis, 21: 349-365, 1984.

____: Quine on underdetermination. In: Perspectives on Quine. R. Barrett and R. Gibson (eds.), pp. 38-52. Cambridge, Mass.: Basil Blackwell, 1990.

____: Quine, underdetermination, and skepticism. Journal of Philosophy, 90: 331-358, 1993.

____: Underdetermination of physical theory. In: The Cambridge Companion to Quine, pp. 91-114. Cambridge: Cambridge University Press, 2004.

Bird, A.: Kuhn's wrong turning. Studies in History and Philosophy of Science, 33: 443-463, 2002.

____: Kuhn, nominalism, and empiricism. Philosophy of Science, 70: 690-719, 2003.

Bonk, T.: Newtonian gravity, quantum discontinuity, and the determination of theory by evidence. Synthese, 112: 53-73, 1997.

Borges, J.L.: El idioma analítico de John Wilkins. In: http://crockford.com/wrrrld/wilkins.html

____: Averroës' search. In: The Aleph and Other Stories. New York: Penguin Books, 2004.

Boyd, R.N.: Realism, underdetermination, and a causal theory of evidence. Noûs, 7: 1-12, 1973.

____: Scientific realism and naturalistic epistemology. PSA: Proceedings of the Biennual Meeting of the Philosophy of Science Association, Volume Two: Symposia and Invited Papers, 1980: 613-662, 1980.

____: The current status of scientific realism. In: Scientific Realism. J. Leplin (ed.), pp. 41-82. Berkeley, Calif.: University of California Press, 1984.

Carnap, R.: Testability and meaning. Philosophy of Science, 3: 419-471, 1936.

____: Testability and meaning – Continued. Philosophy of Science, 4: 1-40, 1937.

____: The Logical Structure of the World. Berkeley, Calif.: University of California Press, 1967.

____: The Logical Syntax of Language. Chicago: Open Court, 2002 [1st English ed. 1937]

Cavell, S.: The Claim of Reason: Wittgenstein, Skepticism, Morality, and Tragedy. Oxford: Oxford University Press, 1979.

Chisholm, R.: The First Person: An Essay on Reference and Intentionality. Minneapolis: University of Minnesota Press, 1982.

Chomsky, N.: Quine's empirical assumptions. In: Words and Objections: Essays on the Work of W.V. Quine. D. Davidson and J. Hintikka (eds.), pp. 53-68. Dordrecht: D. Reidel, 1969.

Churchland, P.M. and Hooker, C.A. (eds.): Images of Science: Essays on Realism and Empiricism, with a Reply from Bas C. van Fraassen. Chicago: University of Chicago Press, 1985.

Clendinnen, F.J.: Realism and the underdetermination of theory. Synthese, 81: 63-90, 1989.

Coffa, J.A.: The Semantic Tradition from Kant to Carnap: To the Vienna Station. Cambridge: Cambridge University Press, 1991.

Cordero, A.: Realism and underdetermination: some clues from the practices-up. Philosophy of Science, 68 (Proceedings): S301-S312, 2001.

Craig, W.: Replacement of auxiliary expressions. The Philosophical Review, 65: 38-55, 1956.

Davidson, D.: On the very idea of a conceptual scheme. In: Inquiries into Truth and Interpretation. pp. 183-198. Oxford: Oxford University Press, 1984.

____: Radical interpretation. In: Inquiries into Truth and Interpretation. pp. 125-139. Oxford: Oxford University Press, 1984.

____: The structure and content of truth. Journal of Philosophy, 87: 279-328, 1990.

____: What is Quine's view of truth? Inquiry, 37: 347-340, 1994.

Dennett, D.: Intentional systems. Journal of Philosophy, 68: 87-103, 1971.

Diamond, C.: Losing your concepts. Ethics, 98: 255-277, 1988.

____: The Realistic Spirit. Cambridge, MA: MIT Press, 1991.

De Rosa, R. and Lepore, E.: Quine's meaning holisms. In: The Cambridge Companion to Quine. R.F Gibson, Jr. (ed.), pp. 65-90. Cambridge: Cambridge University Press, 2004.

Devitt, M.: Underdetermination and realism. Philosophical Issues, 12 (Realism and Relativism): 26-50, 2002.

Douven, I.: The anti-realist argument for underdetermination. Philosophical Quarterly, 50: 371-375, 2000.

____: Evidence, explanation, and the empirical status of scientific realism. Erkenntnis, 63: 253-391, 2005.

Douven, I. and Horsten, L.: Earman on underdetermination and empirical indistinguishability. Erkenntnis, 49: 303-320, 1998.

Dreben, B.: Putnam, Quine – and the facts. Philosophical Topics, 20: 293-313, 1992.

____: *In mediis rebus*. Inquiry, 37: 441-447, 1994.

Duhem, P.: The Aim and Structure of Physical Theory. 2nd ed. and transl. by P. Wiener (Originally published as La Théorie Physique: Son Objet et sa Structure. Paris, Marcel Rivièra, 1914) Princeton: Princeton University Press, 1954.

Dummett, M.: Frege: Philosophy of Language. 2nd ed. Cambridge, Mass.: Harvard University Press, 1981 [1st ed. 1973].

Earman, J.: Underdetermination, realism and reason. Midwest Studies in Philosophy, 18 (Philosophy of Science): 19-38, 1993.

Einstein, A.: On the method of theoretical physics. Philosophy of Science, 1: 163-169, 1934.

Enç, B.: Reference of theoretical terms. Noûs, 10: 261-282, 1976.

English, J.: Underdetermination: Craig and Ramsey. Journal of Philosophy, 70: 453-462, 1973.

Feyerabend, P.: Explanation, reduction, and empiricism. Minnesota Studies in the Philosophy of Science, 3 (Scientific Explanation, Space and Time):27-97, 1962.

____: Science in a Free Society. London: New Left Books, 1978.

____: Putnam on incommensurability. British Journal for the Philosophy of Science, 38: 75-81, 1987.

Field, H.: Science without Numbers. Oxford: Blackwell, 1980.

Fine, A.: And not anti-realist either. Noûs, 18: 51-65, 1984a.

____: The natural ontological attitude. In: Scientific Realism. J. Leplin (ed.), pp. 83-107. Berkeley, Calif.: University of California Press, 1984b.

____: Unnatural attitudes: realist and instrumentalist attachments to science. Mind, 95: 149-179, 1986.

Fodor, J., Lepore, E.: Holism: A Shopper's Guide. Oxford: Blackwell, 1992.

Føllesdal, D.: Indeterminacy of translation and under-determination of the theory of nature. Dialectica, 27: 298-301, 1973.

Friedman, M.: Physicalism and the indeterminacy of translation. Noûs, 9: 353-374, 1975.

____: Theoretical explanation. In: Reduction, Time, and Reality. Richard Healey (ed.), pp. 1-16. Cambridge: Cambridge University Press, 1981.

____: Reconsidering Logical Positivism. Cambridge: Cambridge University Press, 1999.

Gähde, U.: Holism, underdetermination, and the dynamics of empirical theories. Synthese, 130: 69-90, 2002.

Gale, R.: Tensed statements. The Philosophical Quarterly, 12: 53-59, 1962.

____: A reply to Smart, Mayo and Thalberg on "Tensed statements". The Philosophical Quarterly, 13: 351-356, 1963.

Gemes, K.: The indeterminacy thesis reformulated. Journal of Philosophy, 88: 91-108, 1991.

____: Hypothetico-deductivism: the current state of play; the criterion of empirical significance: endgame. Erkenntnis, 49: 1-20, 1998.

____: Hypothetico-deductivism: incomplete but not hopeless. Erkenntnis, 63: 139-147, 2005.

Gibson Jr., R.F.: Are there really two Quines? Erkenntnis, 15: 349-370, 1980.

____: Quine's dilemma. Synthese, 69: 27-39, 1986.

____: Enlightened Empiricism. Tampa: University of South Florida Press, 1988.

____: Quine, Wittgenstein and holism. In: Wittgesntein and Quine. R.L. Arington and H.-J. Glock (eds.), pp. 80-96. London: Routledge, 1996.

____: Translation, physics and facts of the matter. In: The Philosophy of W.V. Quine, 2nd exp. ed. L.E. Hahn and P.A. Schilpp (eds.), pp. 139-154. La Salle, Ill.: Open Court, 1998 [1st ed. 1986].

Giere, R.N.: Scientific realism: old and new problems. Erkenntnis, 63: 149-165, 2005.

Glymour, C.: Theoretical realism and theoretical equivalence. Boston Studies in the Philosophy of Science, 8: 275-288, 1971.

____: Indistinguishable space-times and the fundamental group. In: Minnesota Studies in the Philosophy of Science (Foundations of Space-Time Theories), 8: 50-60. 1977a.

____: The epistemology of geometry. Noûs, 11: 227-251, 1977b.

____: Discussion: hypothetico-deductivism is hopeless. Philosophy of Science, 47: 322-325, 1980.

____: Explanation and realism. In: Scientific Realism. J. Leplin (ed.), pp. 173-192. Berkeley, Calif.: University of California Press, 1984.

Goodman, N.: Fact, Fiction, and Forecast. 4.ed. Cambridge, Mass.: Harvard University Press, 1983 [1st ed. 1954].

Grünbaum, A.: The Duhemian argument. Philosophy of Science, 27: 75-87, 1960.

____: The falsiability of theories: total or partial? A contemporary evaluation of the Duhem-Quine thesis. Synthese, 14: 17-34, 1962.

Hacker, P.M.S.: On Davidson's idea of a conceptual scheme. The Philosophical Quarterly, 46: 289-307, 1996.

Hacking, I.: Representing and Intervening: Introductory Topics in the Philosophy of Natural Science. Cambridge: Cambridge University Press, 1983.

____: Experimentation and scientific realism. In: Scientific Realism. J. Leplin (ed.), pp. 154-172. Berkeley, Calif.: University of California Press, 1984.

____: How inevitable are the results of successful science? Philosophy of Science, 67 (Proceedings): S58-S71, 2000.

Harman, G. Analyticity regained? Noûs, 30: 392-400, 1996.

____: The future of the a priori. In: Philosophy in America at the Turn of the Century (APA Centennial Supplement to the Journal of Philosophical Research). pp. 23-34. Charlottesville, Va.: Philosophy Documentation Center, 2003.

Hart, W.D.: Skolem's promises and paradoxes. Journal of Philosophy, 67: 98-109, 1970.

____: Benacerraf's dilemma. Crítica, 23:87-103, 1991.

Hartshorne, R.: Geometry: Euclid and Beyond. New York: Springer-Verlag, 2000.

Hempel, C.G.: Studies in the logic of confirmation (I.). Mind, 54: 1-26, 1945.

____: Studies in the logic of confirmation (II.). Mind, 54: 97-121, 1945.

Hoefer, C., Rosenberg, A.: Empirical equivalence, underdetermination, and systems of the world. Philosophy of Science, 61: 592-607, 1994.

Horwich, P.: How to choose between empirically indistinguishable theories. Journal of Philosophy, 79: 61-77, 1982.

____: On the nature and norms of theoretical commitment. Philosophy of Science, 58: 1-14, 1991.

Hudson, R.D.: Classical physics and early quantum theory: a legitimate case of theoretical underdetermination. Synthese, 110: 117-256, 1997.

Hughes, R.I.G.: The Structure and Interpretation of Quantum Mechanics. Cambridge, Mass.: Harvard University Press, 1989.

Humphries, B.M.: Indeterminacy of translation and theory. Journal of Philosophy, 67:167-178, 1970.

Hylton, P.: Analyticity and the indeterminacy of translation. Synthese, 52: 167-184, 1982.

____: Translation, meaning, and self-knowledge. Proceedings of the Aristotelian Society, 91: 268-290, 1991.

____: Quine's naturalism. Midwest Studies in Philosophy, 19: 261-283, 1994.

____: Rorty and Quine on scheme and content. Philosophical Topics, 25: 67-86, 1997.

____: Holism and analyticity in Quine's thought. The Harvard Review of Philosophy, 10: 11-26, 2002.

____: Quine on reference and ontology. In: The Cambridge Companion to Quine. R.F Gibson, Jr. (ed.), pp. 115-150. Cambridge: Cambridge University Press, 2004.

____: Quine. (forthcoming)

Jones, W.E.: Underdetermination and the explanation of theory-acceptance: a response to Samir Okasha. International Studies in the Philosophy of Science, 14: 299-304, 2000.

Kirk, R.: Underdetermination of theory and indeterminacy of translation. Analysis, 33: 195-201, 1973.

Kitcher, P.: Theories, theorists and theoretical change. Philosophical Review, 87: 519-547, 1978.

____: The Advancement of Science: Science without Legend, Objectivity without Illusions. New York: Oxford University Press, 1993.

____: Real realism: the Galilean strategy. The Philosophical Review, 110: 151-197, 2001.

Kuhn, T.: The Structure of Scientific Revolutions. 2nd ed. Chicago: University of Chicago Press, 1970 [1st ed. 1962].

____: Dubbing and redubbing: The vulnerability of rigid designation. Minnesota Studies in the Philosophy of Science, 14: 298-318, 1990.

____: Remarks on incommensurability and translation. In: Incommensurability and Translation: Kuhnian Perspectives on Scientific Communication and Theory Change. R.R. Favretti, G. Sandri, and R. Scazzieri (eds.), pp. 33-37.Cheltenham, UK: Edward Elgar, 1999.

____: Commensurability, comparability, communicability. [1983] In: The Road since *Structure*. J. Conant and J. Haugeland (eds.), pp. 33-57. Chicago: University of Chicago Press, 2000a.

____: Metaphor in science. [1979] In: The Road since *Structure*. J. Conant and J. Haugeland (eds.), pp. 196-207. Chicago: University of Chicago Press, 2000b.

____: Reflections on my critics. [1970b] In: The Road since *Structure*. J. Conant and J. Haugeland (eds.), pp. 123-175. Chicago: University of Chicago Press, 2000c.

____: What are scientific revolutions? [1987] In: The Road since *Structure*. J. Conant and J. Haugeland (eds.), pp. 13-32. Chicago: University of Chicago Press, 2000d.

____: Theory change as structure change: comments on the Sneed formalism. [1976] In: The Road since *Structure*. J. Conant and J. Haugeland (eds.), pp. 176-195. Chicago: University of Chicago Press, 2000e.

Kukla, A.: Laudan, Leplin, empirical equivalence, and underdetermination. Analysis, 53: 1-7, 1993.

____: Non-empirical theoretical virtues and the argument from underdetermination. Erkenntnis, 41: 157-170, 1994a.

____: Scientific realism, scientific practice, and the natural ontological attitude. British Journal for the Philosophy of Science, 45: 955-975, 1994b.

____: Studies in Scientific Realism. New York: Oxford University Press, 1998.

____: Theoreticity, underdetermination, and the disregard for bizarre scientific hypotheses. Philosophy of Science, 68: 21-35, 2001.

Ladyman, J.: Understanding Philosophy of Science. London: Routledge, 2002.

Lakatos, I.: Proofs and Refutations. Cambridge: Cambridge University Press, 1976.

Laudan, L.: Grünbaum on the "Duhemian Argument". Philosophy of Science, 32:295-299, 1965.

____: A confutation of convergent realism. Philosophy of Science, 48:19-49, 1981. [Reprinted in: Scientific Realism. J. Leplin (ed.), pp. 218-249. Berkeley, CA: University of California Press, 1984.]

____: Demystifying underdetermination. Minnesota Studies in the Philosophy of Science, 14: 267-297, 1990.

____: Beyond Positivism and Relativism: Theory, Method and Evidence. Boulder, Col.: Westview Press, 1996.

Laudan, L. and Leplin, J.: Empirical equivalence and underdetermination. Journal of Philosophy, 88:449-472, 1991.

Laymon, R.: The path from data to theory. In: Scientific Realism. J. Leplin (ed.), pp. 108-123. Berkeley, Calif.: University of California Press, 1984.

Leplin, J.: Truth and scientific progress. In: Scientific Realism. J. Leplin (ed.), pp. 193-217. Berkeley, Calif.: University of California Press, 1984.

____: Realism and methodological change. PSA: Proceedings of the Biennual Meeting of the Philosophy of Science Association, Volume One: Contributed Papers, 1992: 435-445, 1992.

____: A Novel Defense of Scientific Realism. New York: Oxford University Press, 1997a.

____: The underdetermination of total theories. Erkenntnis, 47: 203-215, 1997b.

____: The epistemic status of auxiliary hypotheses: a reply to Douven. The Philosophical Quarterly, 50: 376-380, 2000

Leplin, J. and Laudan, L.: Determination undeterred: reply to Kukla. Analysis, 53: 8-16, 1993.

Levin, M.: What kind of explanation is truth? In: Scientific Realism. J. Leplin (ed.), pp. 124-139. Berkeley, Calif: University of California Press, 1984.

List, C.: Craig's theorem and the empirical underdetermination thesis reassessed. Disputatio, 7: 28-39, 1999.

MacIntyre, A.: Incommensurability, truth, and the conversation between Confucians and Aristotelians about the virtues. In: Culture and Modernity: East-West Philosophic Perspectives. E. Deutsch (ed.), pp. 104-122. Honolulu: University of Hawaii Press, 1991.

Magnus, P.D.: Underdetermination and the problem of identical rivals. Philosophy of Science, 70: 1256-1264, 2003.

____: Background theories and total science. Philosophy of Science (forthcoming), 2005a.

____: Peirce: underdetermination, agnosticism, and related mistakes. Inquiry, 48: 26-37, 2005b.

____: Hormone research as an exemplar of underdetermination. Studies in History and Philosophy of Biological and Biomedical Sciences, 36: 559-567, 2005c.

____: Reckoning the shape of everything: underdetermination and cosmotopology. British Journal for the Philosophy of Science, 56: 541-577, 2005d.

____: What's new about the new induction? Synthese, 148: 295-301, 2006.

____: Background theories and total science. (forthcoming in Philosophy of Science)

Martens, D.: Demonstratives, descriptions, and knowledge: a critical study of three recent books. Philosophy and Phenomenological Research, 54: 947-963, 1994.

Massimi, M.: What demonstrative induction can do against the threat of underdetermination: Bohr, Heisenberg, and Pauli on Spectroscopic Anomalies (1921-24), Synthese, 140: 243-177, 2004.

McMullin, E.: A case for scientific realism. In: Scientific Realism. J. Leplin (ed.), pp. 8-40. Berkeley, Calif.: University of California Press, 1984.

____: Van Fraassen's unappreciated realism. Philosophy of Science, 70: 455-478, 2003.

Möbius, A.F.: On higher space (1827). In: The Philosophy of Right and Left: Incongruent Counterparts and the Nature of Space. J. Van Cleve and R.E. Frederick (eds.), pp. 39-41. Dordrecht: Kluwer, 1991.

Muller, F.A.: The equivalence myth of quantum mechanics – Part I. Studies in History and Philosophy of Modern Physics, 28(1): 35-61, 1997a.

____: The equivalence myth of quantum mechanics – Part II. Studies in History and Philosophy of Modern Physics, 28(2): 219-247, 1997b.

____: The equivalence myth of quantum mechanics – Addendum. Studies in History and Philosophy of Modern Physics, 30(4): 543-545, 1999.

Newton-Smith, W.: The underdetermination of theory by data. Proceedings of the Aristotelian Society, suppl. vol. 52: 71-91, 1978.

____: Modest realism. PSA: Proceedings of the Biennual Meeting of the Philosophy of Science Association, Volume Two: Symposia and Invited Papers, 1988: 179-189, 1988.

Nola, R.: Fixing the reference of theoretical terms. Philosophy of Science, 47: 505-531, 1980.

Norton, J.: The determination of theory by evidence: the case for quantum discontinuity 1900-1915. Synthese, 97: 1-31, 1993.

____: Science and certainty. Synthese, 99: 3-22, 1994.

____: A little survey of induction. (http://philsci-archive.pitt.edu/archive/00001446/02/Norton.pdf), 2003a.

____: A material theory of induction. Philosophy of Science, 70: 647-670, 2003b.

____: Must evidence underdetermine theory? (http://philsci-archive.pitt.edu/archive/00001257/), 2003c.

Okasha, S.: Laudan and Leplin on empirical equivalence. British Journal for the Philosophy of Science, 48: 251-256, 1997.

____: Holism about meaning and about evidence: in defence of W.V. Quine. Erkenntnis, 52: 39-61, 2000a.

____: The explanation of scientific belief: a response to W.E. Jones. International Studies in the Philosophy of Science, 14: 305-306, 2000b.

____: The underdetermination of theory by data and the "strong programme" in the sociology of knowledge. International Studies in the Philosophy of Science, 14: 283-297, 2000c.

____: Verificationism, realism and skepticism. Erkenntnis, 55: 371-385, 2001.

____: Underdetermination, holism and the theory/data distinction. The Philosophical Quarterly, 52: 303-319, 2002.

Park, S-J.: Hypothetico-deductivism is still hopeless. Erkenntnis, 60: 229-234, 2004.

Peijnenburg, J. and Hünneman, R.: Translations and theories: on the difference between indeterminacy and underdetermination. Ratio (New Series), 14: 18-32, 2001.

Perry, J.: The problem of the essential indexical. Noûs, 13: 3-21, 1979. [Reprinted in: Propositions and Attitudes. N. Salmon and S. Soames (eds.). New York: Oxford University Press, 1988.]

Poincaré, H.: Science and Hypothesis. New York: Dover, 1952 [1905].

Prior, A.: Thank goodness that's over! Philosophy, 34: 12-17, 1959.

Psillos, S.: Tracking the real: through thick and thin. British Journal for the Philosophy of Science, 55: 393-409, 2004.

____: Scientific realism and metaphysics. Ratio (new series), 18: 385-404, 2005

Putnam, H.: Explanation and reference. [1973] In: Mind, Language and Reality: Philosophical Papers, volume 2. Cambridge: Cambridge University Press, 1975.

____: What is 'Realism'? Proceedings of the Aristotelian Society, 75/76: 177-194, 1976.

____: Reason, Truth and History. Cambridge: Cambridge University Press, 1981.

____: What is realism? In: Scientific Realism. J. Leplin (ed.), pp. 140-153. Berkeley, Calif.: University of California Press, 1984.

____: Replies. Philosophical Topics, 20: 347-408, 1992.

Quine, W.V.: Word and Object. Cambridge, Mass.: MIT Press, 1960.

____: Epistemology naturalized. In: Ontological Relativity and Other Essays. pp. 69-90. New York: Columbia University Press, 1969.

____: Reply to Chomsky. In: Words and Objections: Essays on the Work of W.V. Quine. D. Davidson and J. Hintikka (eds.), pp. 302-311. Dordrecht: D. Reidel, 1969.

____: On the reasons for indeterminacy of translation. Journal of Philosophy, 67: 178-183, 1970.

____: The Roots of Reference. La Salle, Ill.: Open Court, 1974.

____: On empirically equivalent systems of the world. Erkenntnis, 9: 313-328, 1975.

____: The nature of natural knowledge. In: Mind and Language. S. Guttenplan (ed.), pp. 67-81. Oxford: Oxford University Press, 1975.

____: A comment on Grünbaum's claim. In: Can Theories be Refuted? Essays on the Duhem-Quine Thesis. (Synthese Library, vol. 81) S.G. Harding (ed.), p. 132. Dordrecht: D. Reidel, 1976.

____: Carnap and logical truth. In: The Ways of Paradox and Other Essays. rev. and enlarged ed. pp. 107-132. Cambridge, Mass.: Harvard University Press, 1976.

____: Posits and reality [1955]. In: The Ways of Paradox and Other Essays, rev. and enlarged ed. pp. 246-254. Cambridge, Mass.: Harvard University Press, 1976.

____: Whither physical objects? Boston Studies in the Philosophy of Science, 39: 303-310, 1976.

____: The scope and language of science [1954]. In: The Ways of Paradox and Other Essays. rev. and enlarged ed. pp. 228-245. Cambridge, Mass.: Harvard University Press, 1976.

____: Facts of the matter. Southwestern Journal of Philosophy, 9: 155-169, 1978.

____: Use and its place in meaning. Erkenntnis, 13: 1-8, 1978.

____: Comments on Newton-Smith. Analysis, 39: 66-67, 1979.

____: From a Logical Point of View. 2nd rev. ed. Cambridge, Mass.: Harvard University Press, 1980 [1st ed. 1953].

____: On what there is. In: From a Logical Point of View. 2nd rev. ed. pp. 1-19. Cambridge, Mass.: Harvard University Press, 1980 [1st ed. 1953].

____: Two dogmas of empiricism [1951]. In: From a Logical Point of View. 2nd rev. ed. pp. 20-46. Cambridge, Mass.: Harvard University Press, 1980.

____: Empirical content. In: Theories and Things. pp. 24-30. Cambridge, Mass.: Harvard University Press, 1981.

____: Five milestones of empiricism. In: Theories and Things. pp. 67-72. Cambridge, Mass.: Harvard University Press, 1981.

____: Goodman's Ways of Worldmaking. In: Theories and Things. pp. 96-99. Cambridge, Mass.: Harvard University Press, 1981.

____: On the very idea of a third dogma. In: Theories and Things. pp. 38-42. Cambridge, Mass.: Harvard University Press, 1981.

____: What price bivalence? In: Theories and Things. pp. 31-37. Cambridge, Mass.: Harvard University Press, 1981.

____: Theories and Things. Cambridge, Mass.: Harvard University Press, 1981.

____: Things and their place in theories. In: Theories and Things. pp. 1-23. Cambridge, Mass.: Harvard University Press, 1981.

____: Relativism and absolutism. The Monist, 67: 293-295, 1984.

____: Indeterminacy of translation again. Journal of Philosophy, 84: 5-10, 1987.

____: Comment on Bergström. In: Perspectives on Quine. R. Barrett and R. Gibson (eds.), pp. 53-54. Cambridge, Mass.: Basil Blackwell, 1990.

____: Lectures on Carnap [at Harvard University, November 8-22, 1934]. In: Dear Carnap, Dear Van: The Quine-Carnap Correspondence and Related Works. R. Creath (ed.), pp. 46-103. Berkeley, Calif..: University of California Press, 1990.

____: Three indeterminacies. In: Perspectives on Quine. R. Barrett and R. Gibson (eds.), pp. 1-16. Cambridge, Mass.: Basil Blackwell, 1990.

____: 'Two Dogmas' in retrospect. Canadian Journal of Philosophy, 21: 265-274, 1991.

____: Commensurability and the alien mind. Common Knowledge, 1: 1-2, 1992.

____: Pursuit of Truth. rev. ed. Cambridge, Mass.: Harvard University Press, 1992 [1st ed. 1990].

____: Structure and nature. Journal of Philosophy, 89: 5-9, 1992.

____: In praise of observation sentences. Journal of Philosophy, 90: 107-116, 1993.

____: Responses. Inquiry, 37: 495-505, 1994.

____: From Stimulus to Science. Cambridge, Mass.: Harvard University Press, 1995.

____: Naturalism; or, living within one's means. Dialectica, 49: 251-261, 1995.

____: Progress on two fronts. Journal of Philosophy, 93: 159-163, 1996.

____: Reply to J.J.C. Smart. In: The Philosophy of W.V. Quine, 2nd exp. ed. L.E. Hahn and P.A. Schilpp (eds.), pp. 516-518. La Salle, Ill.: Open Court, 1998 [1st ed. 1986].

____: Reply to John Woods. In: The Philosophy of W.V. Quine, 2nd exp. ed. L.E. Hahn and P.A. Schilpp (eds.), pp. 726-728. La Salle, Ill.: Open Court, 1998 [1st ed. 1986].

____: Reply to Jules Vuillemin. In: The Philosophy of W.V. Quine, 2nd exp. ed. L.E. Hahn and P.A. Schilpp (eds.), pp. 619-622. La Salle, Ill.: Open Court, 1998 [1st ed. 1986].

____: Reply to Paul A. Roth. In: The Philosophy of W.V. Quine, 2nd exp. ed. L.E. Hahn and P.A. Schilpp (eds.), pp. 459-461. La Salle, Ill.: Open Court, 1998 [1st ed. 1986].

____: Reply to Roger F. Gibson, Jr. In: The Philosophy of W.V. Quine, 2nd exp. ed. L.E. Hahn and P.A. Schilpp (eds.), pp. 155-157. La Salle, Ill.: Open Court, 1998 [1st ed. 1986].

Quine, W.V. and Ullian, J.: The Web of Belief. New York: Random House, 1970.

Reichenbach, H.: The Philosophy of Space and Time. New York: Dover, 1958.

____: The Rise of Scientific Philosophy. Berkeley, Calif.: University of California Press, 1962.

Richardson, A.: Carnap's Construction of the World. Cambridge: Cambridge University Press, 1998.

Rorty, R.: Indeterminacy of translation and truth. Synthese, 23: 443-462, 1972.

Russell, B.: The Philosophy of Logical Atomism. Chicago: Open Court, 1985 [1st ed. 1918].

Sankey, H.: In defence of untranslatability. Australasian Journal of Philosophy, 68: 1-21, 1990.

____: Incommensurability and indeterminacy of translation. Australasian Journal of Philosophy, 69: 219-223, 1991a.

____: Incommensurability, translation and understanding. The Philosophical Quarterly, 41: 414-426, 1991b.

____: Translation failure between theories. Studies in History and Philosophy of Science, 22: 223-236, 1991c.

____: Kuhn's changing concept of incommensurability. British Journal for the Philosophy of Science, 44: 759-774, 1993.

____: Incommensurability: the current state of play. Theoria, 12: 425-445, 1997.

____: Taxonomic incommensurability. International Studies in the Philosophy of Science, 12: 7-16, 1998.

Sarkar, H.: Empirical equivalence and underdetermination. International Studies in the Philosophy of Science, 14: 187-197, 2000.

Scheffler, I.: Science and Subjectivity. Indianopolis: Bobbs-Merrill, 1967.

Schiffer, S.: The basis of reference. Erkenntnis, 13: 171-206, 1978.

Sellars, W.: Inference and meaning. In: Pure Pragmatics and Possible Worlds: The Early Essays of Wilfrid Sellars. J. Sicha (ed.). Reseda, Calif.: Ridgeview Publishing, 1980.

Shapere, D.: Meaning and scientific change. In: Mind and Cosmos: Essays in Contemporary Science and Philosophy (University of Pittsburg Series in the Philosophy of Science, vol. 3). R. G. Colody (ed.), 41-84. Pittsburgh: University of Pittsburgh Press, 1966.

Sklar, L.: Methodological conservatism. The Philosophical Review, 84: 374-400, 1975.

____: Do unborn hypotheses have rights? Pacific Philosophical Quarterly, 62: 17-29, 1981.

____: Theory and Truth. Oxford: Oxford University Press, 2000.

Stalnaker, R.: Indexical belief. Synthese, 49: 129-151, 1981.

Stanford, P.K.: Refusing the devil's bargain: what kind of underdetermination should we take seriously? Philosophy of Science, 68 (Proceedings): S1-S12, 2001.

Stein, S.I.A.: Conteúdo empírico e subdeterminação. Principia, 2: 205-226, 1998.

Tanji, N.: Quine on theory and language. British Journal for the Philosophy of Science, 40: 233-247, 1989.

van Fraassen, B.: The Scientific Image. Oxford: Clarendon Press, 1980.

____: Glymour on evidence and explanation. Minnesota Studies in Philosophy of Science, 10 (Testing Scientific Theories, ed. by J. Earman): 165-176, 1983.

____: On McMullin's appreciation of realism concerning the sciences. Philosophy of Science, 70: 479-492, 2003.

____: Science as representation: flouting the criteria. Philosophy of Science, 71: 794-804, 2004.

Vuillemin, J.: On Duhem's and Quine's theses. In: The Philosophy of W.V. Quine, 2nd exp. ed. L.E. Hahn and P.A. Schilpp (eds.), pp. 595-618. La Salle, Ill.: Open Court, 1998 [1st ed. 1986].

Wang, X.: Taxonomy, truth-value gaps and incommensurability: a reconstruction of Kuhn's taxonomic interpretation of incommensurability. Studies in History and Philosophy of Science, 33: 465-485, 2002.

Wilkes, K.V.: Real People. Oxford: Clarendon, 1988.

Wilson, M.: The observational uniqueness of some theories. Journal of Philosophy, 77: 208-233, 1980.

Wisdom, J.O.: The incommensurability thesis. Philosophical Studies, 25: 299-301, 1974.

Wittgenstein, L.: Philosophical Investigations. G.E.M. Anscombe and R. Rhees (eds.). Oxford: Blackwell, 1958.

____: On Certainty. New York: Harper & Row, 1969.

Worrall, J.: Scientific realism and scientific change. The Philosophical Quarterly, 32: 201-231, 1982.

Yalçin, Ü.D.: Solutions and dissolutions of the underdetermination problem. Noûs, 35: 394-418, 2001.

VITA

NAME: Rogério Passos Severo

EDUCATION: B.A., Philosophy, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil, 1995

M.A., Philosophy, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil, 2000

Ph.D., Philosophy, University of Illinois at Chicago, Chicago, Illinois, 2006

TEACHING EXPERIENCE: Colégio Israelita Brasileiro, Porto Alegre, Brazil, 1995

Faculdade de Direito, Faculdades Integradas do Instituto Ritter dos Reis, Canoas, Brazil; Assistant Professor: Introduction to Philosophy, Philosophy of Law, Logic and Hermeneutics, 1997-2000

Department of Philosophy, University of Illinois at Chicago, Chicago, Illinois; Teaching Assistant: Introduction to Philosophy, Introductory Logic, Introduction to Social and Political Philosophy, Death, Philosophy and Science, 2001-2006

HONORS: CNPq Fellowship, Brasília, 1995-1997

CAPES Fellowship, Brasília, 2000-2004

PROFESSIONAL MEMBERSHIP: American Philosophical Association

Sociedade Kant Brasileira

ABSTRACTS: Severo, R. P.: Contrapartidas incongruentes. Anais do II Congresso Kant, 1997.

Severo, R. P.: Distinguir, reconhecer e identificar contrapartidas. Anais do IX Encontro Nacional de Filosofia (ANPOF), 2000.

PUBLICATIONS: Kant, I. Sobre o primeiro fundamento da distinção de direções no espaço (traslation of the original German text to the Portuguese, introductory note, endnotes). Cadernos de Filosofia Alemã, 2: 61-75, 1997. (Also published online in at http://www.ufrgs.br/kantcongress/sociedadekant/fundamento.pdf)

Höffe, O. O imperativo categórico do direito (translation of the original German text to the Portuguese, in collaboration with Valério Rohden). Studia Kantiana, 1: 203-236, 1998.

Bittner, R. Máximas (translation of the original German text to the Portuguese, in collaboration with Mauro Engelmann). Studia Kantiana, 5: 7-25, 2004.

Severo, R. P.: Three remarks on the interpretation of Kant on incongruent counterparts. Kantian Review, 9: 30-57, 2005.

THEORIES AND REALITY
FIVE ESSAYS ON QUINE AND UNDERDETERMINATION

Rogério Passos Severo, Ph.D.
Department of Philosophy
University of Illinois at Chicago
Chicago, Illinois (2006)

Dissertation Chairperson: Peter W. Hylton, Ph.D.

The thesis of underdetermination has received considerable attention in the philosophical literature of the last few decades, yet little consensus has been reached about how it is to be thought of or formulated. Even its most influential proponent, W.V. Quine, changed his mind several times about how to characterize the thesis, and vacillated about some of its implications. The present dissertation contains a comparative analysis of Quine's views on underdetermination and contemporary debates on the matter.

Five main points are defended in the dissertation: (i) holism and underdetermination are distinct theses; while the former may be thought to lend credence to the latter, it is insufficient to establish it; (ii) Quine's formulation of the thesis is significantly weaker than commentators have thought; (iii) rivalry among empirically equivalent theories is best understood in terms of non-intertranslatability of fundamental terms or predicates; (iv) some recent criticisms of the thesis are empty in that they unduly strengthen the thesis so as to render it untenable; and some recent defenses of the thesis rely on an overly lax view of what counts as a rival theory, or unduly weaken it in such a way that it ends up indistinguishable from radical skepticism; (v) contrary to a common assumption in the recent literature, underdetermination does not have to be thought as incompatible with epistemic realism.