UD is not a DUD

A novel anti-realist account based on underdetermination.

Master's Thesis

Author: K. Schipper BSc [0448184]

Date: 4-06-2012

Instructor: Prof Dr D.G.B.J. Dieks

Faculty of Beta Sciences / Department of Physics and Astronomy

Institute for the History and Foundations of Science (IGG)

TABLE OF CONTENTS

INTRODUCTION

This master's thesis is a continuation of the research I had started when writing my bachelor's thesis. That thesis had a similar idea but the execution was lacking. What was mostly missing was a sophisticated look at the specific arguments raised against underdetermination itself. The beginning of this thesis is for that reason very much like that bachelor's thesis, but the other 70-odd pages that follow should provide plenty of sophistication.

The main aim is unchanged: I want to advocate a form of anti-realism based on underdetermination. However, the details have changed enormously. I have adopted a more sophisticated form of underdetermination, I have looked in detail at arguments against underdetermination and I have made explicit what I think is the basis of knowledge, and what should be the attitude of the scientist.

The thesis is divided in three chapters. Chapter one and two provide most of the content, while chapter three is merely a conclusion with a few remarks about Structural Realism. The first chapter deals with the question "What kind of underdetermination should we take seriously?". I look at proposed forms of underdetermination in modern literature, and then at the two original authors (Quine and Duhem) of the argument in the scientific context, and I compare these views at the end. I also add discussion about Collins' experimenter's regress which I think should be understood as a modern explication of Duhemian argumentation. Together, these form what I think should be regarded as the form of underdetermination to be taken seriously. It will be named DONUD.

In chapter two I will discuss the question "Is there a way in which DONUD is either rendered unproblematic or solved?". I first introduce an argument aimed at dispelling charges of inductive skepticism. Then I will try to dispel a similar argument aimed at discharging a fundamental distinction between the observable and unobservable world. Then I look at arguments that either show underdetermination to be trivial or provide a way to solve the underdetermination problem. I will show how these are either misrepresentations of the fundamental issues involved or argument which have no force when subjected to further analysis.

Throughout the discussion in these two chapters I will construct a view of anti-realism which avoids all the traditional criticism and which is thoroughly based on underdetermination. In that sense I distinguish myself from other anti-realisms, most notably constructive empiricism.

PART I: WHAT IS UNDERDETERMINATION?

The realism debate in philosophy of science

The realism debate in the philosophy of science is a debate about whether or not we are justified in believing that science is capable of telling us what the world really is like. That is to say, science gives us a certain picture of the world, and the question is whether or not we have good reason to believe that this picture is true in the sense that the events described in the picture are as the events are in reality, perhaps approximately so. The stance that science is *capable* of telling us what the events in reality are and that we are justified in believing as –approximately- true what scientific theories tell us in this way, even if we cannot directly observe the entire reality science presents to us, is called scientific realism. Any stance that opposes this view is called scientific anti-realism. Of course, realists allow that a theory actually available to us does not necessarily tell us what the events in reality are, for instance when a theory happens to be wrong. The debate is about what can be achieved, not what is actually achieved.

There have been various forms and degrees of realism, even outside the philosophy of science, as there have been various forms and degrees of anti-realism. In this thesis whenever I refer to realism or anti-realism, I mean scientific realism or scientific anti-realism as given above. It is important to stress that the part of the realism debate I want to focus on is about the reality of unobservables, be they entities or events or in some exotic cases of realism the (mathematical) relations between these. In this I follow the course of the debate as it developed during the 20th century. Even if realism has always been the more popular position, I intend to show that anti-realism is just as strong, if not stronger, and that the issue is still unresolved. This is in spite of the prevailing attitude of scientists themselves, of whom most seem to have a realist attitude.

Note that realists and anti-realists usually agree on the reality of observable entities, such as tables and chairs being real, your car being stolen and you heading to the police station being real –one would call this a fact if true- and so on. But I want to raise the question whether or not science actually has the power to make claims about the microscopic nature of the universe, such as the origin of life on our planet and the existence of genes, or even the macroscopic nature of the universe as far as we have only indirect evidence about it, such as the existence of planets in other star systems. In the discussion I primarily use examples based on issues in physics; that is just because it is the field I am familiar with. The discussion itself extends itself across all of science as long as it involves indirect evidence, which requires a relation between observation and theory.

The focus on the unobservable world is traditionally the crux of the problem. Science wouldn't be so hotly debated, attacked and defended, particularly outside of the field of philosophy of science, if it would only make claims about things everyone can experience for themselves; to be frank that would be wholly unspectacular. It is rather when science gains this almost mythical power of telling us about the properties of invisible entities, or telling us what is going on at scales so tiny, and I add so large, as to -possibly- go beyond our comprehension, that controversy arises. It should be no surprise this has happened and is happening within the philosophy of science as well.

We find skepticism about claims going beyond what is directly observable throughout history. Mostly this scepticism is part of -what is now interpreted- as a debate about the *source* of our knowledge, be that divine inspiration, sense-experience or ratio. Whatever the stance in this debate, everyone seems to have accepted the use of purely deductive logic since Aristotle. In Aristotle, we find an explicit form of deductive logic in the 'syllogism':

"A syllogism is discourse in which, certain things being stated, something other than what is stated follows of necessity from their being so. I mean by the last phrase that they produce the consequence, and by this, that no further term is required from without in order to make the consequence necessary."[1]

An easy example is the following: if an A is a B, and every B is a C, then necessarily that A is a C. While Aristotle calls this a syllogism and the term syllogism has come to stand for the specific structure of the premise-deduction-conclusion argument, the above cited passage describes something close to our understanding of deduction: reasoning in which we merely explicate what is contained in the premises. There are some subtle differences between what Aristotle has written and our modern understanding of deduction, but these are not relevant to what I want to discuss here. In science, a primary use of deduction has been the inference from universal statements to particular statements. So we infer from the universal second law of Newton that this particular cannonball with mass m when subjected to a force F will have acceleration a.

Inference in the opposite direction is sometimes called induction; induction goes from particulars to universal statements. But there are many different ways to infer from particulars to universals, and the term 'induction' is used for different types of inference. In this thesis I want to be explicit about what I mean with terms like 'induction'. I think it is unnecessarily confusing to use the word as a generalized term for all forms of non-deductive inference. For instance, there is already a difference between enumerative induction and Bayesianism[2], both 'inductive' forms of inference in the sense that they are not deductive forms

[1] Aristotle a
[2] These two forms of inference will be introduced below.

of inference. Also note that deduction need not necessarily flow from universals to particulars either as it can be from particulars to particulars, but never from particulars to universals. There is an exception if the group of particulars is exhausted; we *can* deductively infer from observing several *A's that are B* that *all A's are B* if and only if we happen to have observed *all* A's.

As I mentioned above, everyone seems to accept use of deductive logic, and this should not surprise anyone. Deduction does not really add any new information, though we might come to conclusions which we were not aware of. The information in the conclusion is already contained within the premises. In knowledge claims, the use of deduction can be divided in empirical use and non-empirical use. In the former case, we deduce from observed phenomena and this is usually not problematic. If we can trust our senses or our experience of the phenomena, we can safely deduce from observed facts. In the latter case, we deduce from some non-empirical statement, be that an assumption or a hypothesis. In this case deduction itself is not problematic either; the validity of the conclusion depends on the validity of the non-empirical statements. In this regard it is no different from empirical deduction, but the validity of the premises is a priori less of a problem if the premises have empirical verification.

A crucial question remains: what about the validity of non-empirical statements? If it is the case that we cannot use deduction from observed facts to confirm their validity, and I will argue that this is indeed the case, how else can we do this? We connect this to the realism debate in science by asking ourselves how we can infer from observed facts to statements beyond observation. We cannot deduce from observations statements beyond observation, as deduction only explicates what information is in the observations themselves. In the context of the difference between deduction and induction as introduced above: Our observations are always particulars, so we cannot use deduction to move from these particulars to a universal statement, which is generally supposed to be an aim of science. How can science provide us with justifiable knowledge beyond the observable at all? There are two basic answers to this question, and both have faced their share of criticism. The first is to allow more than deduction, such as specific forms of induction. The second is to allow the use of non-empirical assumptions.

The latter position usually falls back on some kind of use of human reasoning. Using the mind, our *ratio*, we could find knowledge that is for instance intuitively clear or self-evident. The position that one source of knowledge originates in our own faculty of thought has been called 'rationalism', though many different kinds of rationalism exist. The position that the foundation of our knowledge should be information gained by our senses is called 'empiricism', though some positions usually considered empiricism allow degrees of rationalism. Aristotle, for instance, can be regarded as an empiricist even though he allows the use of a form of intuition (self-evidence) to help formulate the axioms of his system of knowledge:

"We conclude that these states of knowledge are neither innate in a determinate form, nor developed from other higher states of knowledge, but from sense-perception. It is like a rout in battle stopped by first one man making a stand and then another, until the original formation has been restored. The soul is so constituted as to be capable of this process.

When one of a number of logically indiscriminable particulars has made a stand, the earliest universal is present in the soul: for though the act of sense-perception is of the particular, its content is universal - is man, for example, not the man Callias. A fresh stand is made among these rudimentary universals, and the process does not cease until the indivisible concepts, the true universals, are established…"[3]

So for Aristotle our soul has an innate capacity to formulate universals out of particulars. But there has been criticism that this kind of process is not very trustworthy at all. What have we really witnessed? Surely it was not the universal *itself*. Aristotle thinks as is shown in the citation above that our soul is capable of 'seeing' universals, in a way. But what if we do not want to go along with such an appeal to rationality? How can the experience of particulars bring us to the universal then? D. Hume has formulated this problem sharply:

"As our senses shew us in one instance two bodies, or motions, or qualities in certain relations of success and contiguity; so our memory presents us only with a multitude of instances, wherein we always find like bodies, motions, or qualities in like relations. From the mere repetition of any past impression, even to infinity, there never will arise any new original idea, such as that of a necessary connexion; and the number of impressions has in this case no more effect than if we confined ourselves to one only." [4]

That is, it does not matter how many times we observe a certain event, the events do not themselves imply a 'connection', a relation. Hume is sceptical about our ability to go beyond the senses:

"[…]it is impossible for us to satisfy ourselves by our reason, why we should extend that experience beyond those particular instances, which have fallen under our observation."[5]

Now we see how the problem is about what goes on beyond the reach of our senses and how we should gain knowledge beyond this reach. For Hume even the simple inductive inference of extending repeating events into the future is going beyond sense experience, and thus is open to criticism. In Hume we find

---

[3] Aristotle b
[4] Hume
[5] Hume

the foundation of a lot of the criticism that has been aimed at rationalism and realism. If we cannot even rely on induction as a tool of generating knowledge as Hume claims, how can we gain knowledge about anything at all, beyond the particulars currently clear to our senses? Can we move from particulars to anything else at all? Can we even extend those particulars to the future?

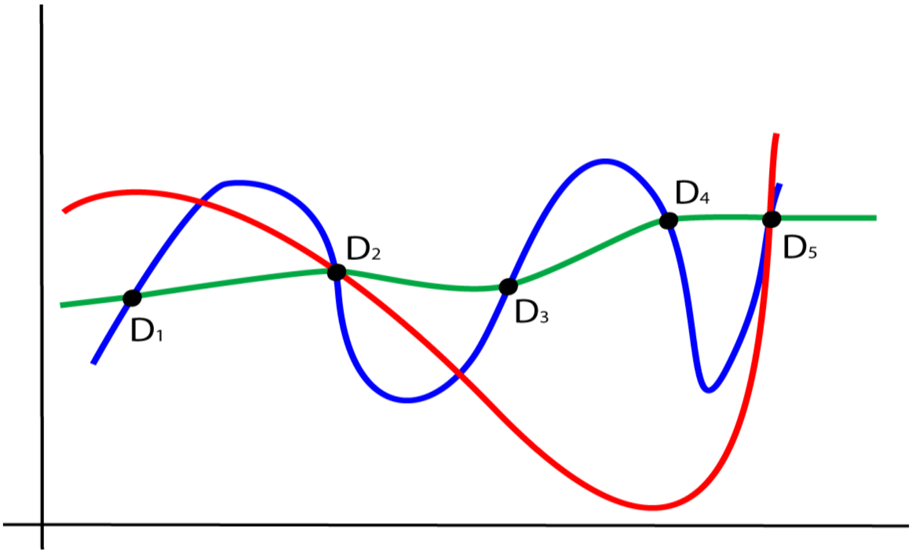
I must point out, to be historically accurate, that philosophers like Hume and Aristotle were not themselves concerned with the idea of *science*. Rather, they were concerned with problems regarding *knowledge*, and it is *our* reflection that we see similarities in arguments and problems presented by these historical figures rather than that they were dealing with the same issues. The debate about *scientific* realism, which includes problems discussed in pre-scientific times such as induction, underdetermination, confirmation and unobservables, is –mostly- a debate of the 20th century and it continues presently. The specific questions discussed here are mostly about science, and not about knowledge in general, though of course questions about the latter automatically apply to the former as science claims to be a form of knowledge, hence the reflection on ancient discussions. What can be said is that many of the problems Hume and other philosophers have pointed out remain points of discussion within philosophy of science. What Hume has pointed out is now debated as the "problem of induction" and remains a focal point within the realism debate even though the scope of the realism debate has increased to include more subtle arguments.

The argument of underdetermination

One important argument in the realism debate is the argument of underdetermination. In the philosophy of science underdetermination is used in the context of the relation between theory and evidence –usually in the form of experiment. Data comes in from certain devices or from experimenters in processed form. If we accept the incoming data as is -which is not an obvious thing to do and usually only the case after a long period of subjugating the available data to tests- this data either confirms or disconfirms a theoretical statement. Sometimes this is a constructive relation, when evidence is used to construct a theory, and sometimes this is an evaluating relation, when evidence is used to test a theory, but neither excludes the other; indeed, in practice the difference is not so easy to make. In the philosophy of science the discussion has mostly been about the latter relation because the focus is on justification. H. Reichenbach coined the terms 'context of discovery' and 'context of justification' to reflect this, but I'm not convinced 'justification' is the proper term, hence my use of 'evaluating'.

Underdetermination steps in when we want to move from the evidence to some generalized form which entails the evidence. The problem of underdetermination is just that the evidence itself does not logically lead to a generalized form. The evidence underdetermines the choice of generalization in the sense that there are always multiple options available logically. In the relation between theory and evidence, this means that evidence does not flow naturally towards theory, that the evidence does not in itself determine or construct theory, and that there must be an additional element which makes that flow possible.

There is a canonical representation of underdetermination as 'curve-fitting'. Imagine five points on a graph which represented the available data D1, D2, D3, D4 and D5 (see the figure below). Our job is to plot a curve through the points that would be the 'best fit'. How do we go about to determine what is the best fit? Suppose that we know points D2 and D5, then all three curves of the figure would fit. The curve is underdetermined by points D2 and D5. Suppose we now learn D1, D3 and D4, so that rules out the red curve but still leaves the blue and green curve available. The points still underdetermine the curve even though our amount of curves available is lower. In fact, one can imagine an infinity of curves that are possible to fit through D1 to D5.[6]



The idea is that this mathematical representation can be expanded to apply to the relation of evidence and theory in science. This problem should provide plenty of ammunition to empirically minded anti-realists. If evidence does not specifically point to a single theory, then there are multiple theories available given a single body of evidence. But isn't this just to say that we can't be sure about which theory is the correct one? Or in terms of realism and truth, which theory is the proper representation of the events as they are in reality? Thus, a possible anti-realist stance, thoroughly empiricist in nature, is that because of

[6] Lipton mentions this in passing, see Lipton 1998 p. 412

underdetermination we cannot conclude anything about the truth of a theory beyond observable events, because the truth of a theory on the level beyond observable events is left underdetermined by the same observable events.

The actual discussion on underdetermination as it has taken place in the last century is largely based on the works of P. Duhem and W.V.O. Quine. It is illustrative of their impact that the 'thesis of underdetermination' is often called 'the Duhem-Quine thesis', though it is over-simplifying matters to speak of a Duhem-Quine thesis, as we shall see. Many arguments against underdetermination made by realists are aimed at Duhem or Quine. In the discussion we rarely find an account of underdetermination in its primal form as sketched above. Articles on underdetermination usually concern two *logical consequences* of underdetermination; the *problem of auxiliary hypotheses* or the *problem of empirically equivalent theories*. For instance, Laudan and Leplin state in their now almost canonical rebuttal of empirical equivalence:

"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. Secondly, by the 1940s and 1950s, it was thought that – in large part because of empirical equivalence – theory choice was radically underdetermined by any conceivable evidence."[7]

In another article Laudan gives us another consequence of underdetermination; the use of auxiliary hypotheses to 'rescue' our theories from seemingly falsifying instances. He presents a form of underdetermination called Quinean underdetermination (QUD*)[8] as:

'Any theory can be reconciled with any recalcitrant evidence by making suitable adjustments in our other assumptions about nature.'[9]

This characterization of underdetermination is largely based on the works of Quine, to whom we will turn shortly.

What I want to make clear is that what is commonly referred to as the 'underdetermination thesis' in fact consists of two slightly different arguments; the 'problem of auxiliaries' and 'the problem of empirical equivalence'. I want to note that they are not logically equivalent, and not necessarily logically equivalent to underdetermination itself, and that I want to avoid confusion about the terms. If we regard 'underdetermination' as the term describing all problems of the kind in which there is not enough information to determine a solution, then there are different kinds of underdetermination. At the end of this chapter, I wish to present a clear view of what I mean by 'underdetermination', and I want to avoid the confusion due to using the term 'underdetermination' while referring

[7] Laudan and Leplin 1991, p. 449
[8] The asterisk is there to avoid confusing QUD* as QUD as I will present later.
[9] Laudan 1998 p. 328

to distinct arguments involving underdetermination. I wish to present a specific underdetermination argument as the *prime contender*.

K. Stanford does recognize explicitly the difference between the problem of auxiliaries and the problem of equivalence and calls these two problems 'holist underdetermination' and 'contrastive underdetermination' respectively:

"*Holist underdetermination* […] *arises whenever our inability to test hypotheses in isolation leaves us underdetermined in our response to a failed prediction or some other piece of disconfirming evidence: that is, because hypotheses have empirical implications or consequences only when conjoined with other hypotheses and/or background beliefs about the world, a failed prediction or falsified empirical consequence typically leaves open to us the possibility of blaming and abandoning one of these background beliefs and/or 'auxiliary' hypotheses rather than the hypothesis we set out to test in the first place. But contrastive underdetermination* […] *involves the quite different possibility that for any body of evidence confirming a theory, there might well be other theories that are also well confirmed by that very same body of evidence."*[10]

A more technical approach to underdetermination is found in Earman's book on Bayesianism. He discusses J. Dorling's Bayesian solution to the problem of auxiliaries. If we suppose that a theory T consists of core hypotheses T_1 and auxiliary assumptions T_2 so that T_1 & $T_2 \vDash \neg E$ and the result of an investigation is E, then, given certain assumptions about the probabilities of T_1, T_2, E and their prior probabilities and likelihoods, we can show how the evidence falls more squarely on T_2 than T_1. This would suggest that through the use of Bayesian rules we can infer that, in contrast with the problem of auxiliaries, the auxiliaries are more likely to be the culprit here than the core hypotheses. Thus, we have some way out of the problem of auxiliaries; we cannot always adjust our theory such that it is compatible with the evidence without violating the laws of probability according to Bayesian analysis. This form of underdetermination is presented in the form of the problem of auxiliaries.[11]

This is so far a quick look at the place of underdetermination in the realism debate and how it has been presented. To be clear, let's sum up what we have so far. Underdetermination is a problem in the relation of theory and evidence. It is still unclear what exactly it is, but this will hopefully be solved at the end of this chapter. We have so far two corollaries of underdetermination which I will generalize as much as possible:

Holist underdetermination or **the problem of auxiliaries**: For any theory T, if an hypothesis H is a logical consequence of T, and we predict evidence E by adding auxiliary hypotheses A to H, if it is the case that from experiment follows that E is false, then we can always adjust A or T so that it follows from H that E

[10] Stanford 2009
[11] Earman 1992, p. 83

is false. Likewise, if we predict evidence ¬E and we find E, we can adjust so that E is true. There is a stronger form with the addition: we can always adjust A in such a way that this adjustment is scientifically acceptable.

Contrastive underdetermination or **the problem of empirical equivalence**: For any set of evidence $S_{\mathcal{E}}$, we can find theories T and T' so that $S_{\mathcal{E}}$ cannot be of help in determining which of T and T' is true. There is a stronger form with the addition: $S_{\mathcal{E}}$ cannot be of help for the purposes of scientific choice either.

But why should either of these corollaries of underdetermination hold? In order to answer this question, we must now turn to Duhem and Quine. As mentioned before the problem of underdetermination is sometimes over-simplified as the "Duhem-Quine thesis", but a careful read of both reveals that they have different conceptions of underdetermination. D. Gillies has written an article about the particular differences between the Duhem and Quine theses[12], and his account of three main differences should be enough to convince anyone that the Quine and Duhem theses deserve a separate account.

Duhemian underdetermination

The arguments involving underdetermination are often called the "Duhem-Quine thesis" after P. Duhem and W.V.O. Quine, two philosophers well known for their anti-realism based on a form of underdetermination. But as mentioned Quine and Duhem did not claim the same thing, and let's not forget there are a few decades between their respective publications, putting their arguments in different contexts. I will discuss and consider both accounts separately, and explicate the important arguments for my thesis while trying to stay faithful to their claims. I will also present Harry Collins' book 'Changing Order' as a *Gruesome New Account of Underdetermination*, and compare it with Duhem's article to show why I think his book is still the *magnum opus* in the literature on underdetermination. For now let us focus on this first account of underdetermination in the philosophy of science.

In 1906 the first edition of "The Aim and Structure of Physical Theory" was released. In it Duhem wants to "offer a simple logical analysis of the method by which physical science makes progress."[13] Only two chapters at the end of the book really deal with what we now call the "Duhem thesis", the first of which is the most famous. Notably, it has been extracted from the book and presented as a stand-alone article in Curd and Cover's "Philosophy of Science: the Central Issues". Duhem's thesis has become known as 'physical holism', since his main point is that the physical scientist is always employing a multitude or, more aptly, an entanglement of theories. The scientist is not only employing the

[12] Gillies 1998, p. 302-319
[13] Duhem 1982, p.3

hypothesis to be researched when focusing his work on a certain hypothesis or another. He is also using a whole network of hypotheses and theories. For instance, he relies on scientific instruments used, and thus relies on the theories that govern those instruments. But the scientist also relies on the theoretical framework surrounding his work to guide his decisions. Duhem puts it like this:

"We have seen that in the mind of the physicist there are constantly present two sorts of apparatus: one is the concrete [...] the other is the schematic and abstract [...] which theory substitutes for the concrete [...] and on which the physicist does his reasoning. For these two ideas are indissolubly connected in his intelligence, and each necessarily calls on the other."[14]

The scientist cannot isolate one hypothesis from a group of others connected with the one hypothesis in the context of a certain experiment, and as a result he cannot test any hypothesis in isolation. For a test of a hypothesis must be devised and this process itself is guided by theory, whatever theory the physicist deems appropriate. One cannot measure an electron without having an idea of what properties apply to an electron. Since we cannot observe the electron directly we have to devise instruments which allow us interaction with it using those properties, and these instruments are in need of interpretation of their own. Now we see how quickly the amount of theory involved increases just looking at a single hypothesis involved in a single experiment, since the instrument comes with a whole framework of knowledge on how to operate it and this in turn creates other dependencies on theoretical knowledge.

As a result of this an experiment can never condemn a single hypothesis but only a group of hypotheses as a whole, namely the entire group that is invoked in the experiment. When an experiment reveals an inconsistency, all the scientist really has learned is that at least one of the hypotheses involved in the experiment is wrong, but not *which* of these it is.

"Physics is not a machine which lets itself be taken apart; we cannot try each piece in isolation and, in order to adjust it, wait until its solidity has been carefully checked."[15]

Duhem adds that physics must be taken as a whole and that even the 'most remote' parts are in use. As a point of criticism, this seems to be taking it too far. It is certainly plausible that for a single experiment, or for that matter any activity in physics, a group of hypotheses is used directly or indirectly. It is a whole other matter whether or not *all* physical hypotheses -currently accepted- come into play and this would require more support than what Duhem gives in his article.

It is in this light interesting to read Gillies' attempt to separate Duhem and Quine. One of his points is that Duhem, in contrast with Quine, places limits on

[14] Duhem 1982, p. 183
[15] Duhem 1982, p. 187

the size of the group of hypotheses actually in play in the situation of which Duhem's thesis is typical.[16] His criticism of Quine is that Quine considers the problem of underdetermination to be valid for the entirety of knowledge at once, while Gillies thinks that depending on the subject not every empirical claim is used. He uses the example that while Adams and Leverrier used an extensive amount of hypotheses to deduce certain conclusions about the orbit of Uranus, it seems absurd to assume that a hypothesis like 'bees collect nectar from flowers in order to make honey' is involved in this deduction, at least Adams and Leverrier did not mention this hypothesis. So why does he raise this point for Quine, while Duhem makes a similar claim? While Quine's statement is obviously stronger, because it is about *all* knowledge, Duhem still talks about the *entirety* of physics. I see no a priori reason Duhem's claim should be beyond doubt if Quine's isn't.

Duhem's further analysis in this chapter consists of consequences of what we shall now call his physical holism. He points out how this holism implies that several kinds of physical methodologies cannot be held as justifiable methods of physical science. Eliminative induction, Newtonian induction, reduction to facts of experience; all these methodologies rely in one way or another on the decisiveness of experiment, which according to the physical holist view is an erroneous assumption. Experiment itself relies fundamentally on multiple hypotheses so to say with certainty that one of the hypotheses is the culprit and not one of the others seems impossible. Experiment can only tell you that *something* is wrong, not *what* exactly. I think that in this view experiment cannot tell you logically that something is right either, since it is entirely possible for the experiment to rely on hypotheses some of which are not true, yet lead to the expected result by virtue of their cooperation in removing the error in the end result.

Far worse, it is according to Duhem possible to have statements which have *no physical meaning.*[17] What does Duhem mean with 'physical meaning'? It seems clear from the examples he gives that he means that no experiment can directly refute or verify these principles, or that in principle it is possible to defend such principles from any criticism. I will use an example that is not used by Duhem but which I am more familiar with. Consider the case of Einstein vs. Lorentz, and the principle of the absolute frame of rest, which for Lorenz was the frame in which the ether is at rest. No experiment decided the truth or falsehood of the existence of an absolute frame of rest and it was Einstein who made the point that all inertial frames are equivalent from the point of view of the physical equations. Thus, even if an absolute frame of rest would exist, we would not be able to detect it because it would be physically equivalent to other inertial

[16] Gillies 1998, p.313-314
[17] Duhem 1982, p. 216

frames. In this way it becomes impossible to defend or reject the existence of a frame of absolute rest based on experiment. The frame of absolute rest plays the role of a statement which has no physical meaning.

Another example can show how sometimes we can even *have a choice* between statements which in themselves have no physical meaning. There is a degree of convention involved. Consider the Euclidean and non-Euclidean geometries, and we ask ourselves the question "which of these describes the true geometry of the universe?". It is proven in mathematics that these geometries are consistent and that if one is consistent, the others necessarily are too, and finally that each can be embedded in the other. So, non-Euclidean geometries can be geometries about special surfaces in Euclidean geometry and vice versa. For any experiment which might seem to conclude in favour of one of the geometries over the other, we could always argue that what we were actually measuring was a special surface in the *true* geometry, provided that we introduce a universal force which distorts our measuring rods in such a way as to produce the appearance of the special surface. In fact, we can freely choose one of the geometries as the prime geometry and the other geometries as about special surfaces in the prime geometry and still have a consistent picture which agrees with the same observable facts, provided we introduce those universal forces.

I think that Duhem had this type of statement in mind when he proposed that some statements have no physical meaning. This seems to me similar to the problem of auxiliaries that we have found above, but with one difference. Duhem still thinks there is a possibility for experimental refutation to some degree, even for those statements which have no physical meaning. According to his physical holism, it is only possible to refute an entire system of hypotheses, including these statements which have no physical meaning by themselves. But precisely this fact implies that statements which have no physical meaning by themselves *are tested in the same way as any other statement.* They are in this regard no different from statements which he would consider to have some physical meaning. Interestingly, does this mean that, taken by themselves, *all* statements have no physical meaning in this view? Duhem has a way out of this problem in the form of *le bon sense* of the scientist.

To reiterate, the problem of auxiliaries is about finding auxiliary hypotheses to safeguard a set of core hypotheses. The problem is that you can always find auxiliaries to protect the core of a theory if you are so inclined. In the first paragraph of the last section of this chapter Duhem admits the possibility of the existence of auxiliaries in this way. He also admits that someone who is differently inclined might change the core of the theory involved in the light of disconfirming evidence. There is no logical answer to this problem, simply because experiment does not adjudicate this process. But here Duhem has an interesting twist to the story:

"Pure logic is not the only rule for our judgements; certain opinions which do not fall under the hammer of the principle of contradiction are in any case perfectly unreasonable. These motives which do not proceed from logic and yet direct our choices, these "reasons which reason does not know" and which speak to the ample "mind of finesse" but not to the "geometric mind," constitute what is appropriately called good sense."[18]

What does Duhem mean? What does good sense constitute? We shall see throughout this thesis that although the bon sense is vague, the one point it makes which is shared by almost all philosophers is that something outside pure deductive logic is used. I must add that philosophers of science since Kuhn have tried to point out the 'rational' in this extralogical process, more on that in part II. Kuhn himself proposed *values of science* such as consistency and simplicity[19] and Laudan has suggested so-called ampliative rules of inference which guide this process.[20] More on this in the third chapter, for now we will return to what Duhem has to say about the extralogical process which guides judgements about hypotheses. Duhem dedicates the next chapter in his book, some 50 pages, to expound on this idea of good sense. Let us go through them and note any statements contributing to our discussion.

We can readily note that Duhem feels it's important to look for support in the history of science. His approach, like in some parts of the previous chapter of his book, is to give an account of historical events concerning some scientific theory and to show subsequently how his claims are true for these accounts. In this case what constitutes good sense, and how the physicist uses and gains this good sense is to be found in historical accounts. Indeed, in the last part of this chapter his claim is that the historical method is *the* method by which a physicist can gain good knowledge about what good sense and consequently good science constitutes.

Good sense seems to be part of a process, as Duhem states that 'hypotheses are not the product of a sudden creation, but the result of a progressive evolution' and that 'history shows us that no physical theory has ever been created out of whole cloth.'[21] Hypotheses come about through a slow and gradual process taking place in the scientists' mind, in which a constant reflection takes place on experimental results and these result in the formation of an idea which turns into a hypothesis only at the end of the long and arduous process. Logic does not guide this process and the physicist has a large degree of freedom in this stage of hypothesis formation, but at the end when the process is

[18] Duhem 1982, p. 217
[19] Kuhn 1998, p. 103
[20] Laudan 1998, p. 336
[21] Duhem 1982, p.220-221

completed, the hypothesis must be made consistent with existing knowledge. In the last part of this chapter Duhem makes clear that the only method which is capable of elaborating what is taking place during the time of formation, and from which we gain understanding of the various details of the formation of the current theories which in turn can guide us in our process of creating scientific theories, is the historical method.

Another interesting point Duhem raises in this discussion is that there is a remarkable difference between the phases when theory is being constructed and theory having reached its complete development.[22] When the theory is being constructed, a set of core principles is set down and these are defended at all cost, even in the face of seeming experimental refutation. If we follow Duhem's earlier posited physical holism this would indeed be possible. But when the theory is completed then it must face the tribunal of experimental evidence. This is because only the *conclusions* of theory bear on experimental fact, after all only the conclusions 'are offered as an image of reality;'[23] the postulates only serve as the foundation upon which the conclusions are built. What is remarkable here, and in the paragraph above about the formation of hypotheses, is that Duhem seems to point to the *practice of science*, where theories are being constructed in the lab or in the theoretician's office.

This is a striking difference with normative philosophy of science which is about static relations between experimental facts, which by themselves are considered unproblematic, and theory. Nothing about *the process* of construction of the experimental fact and the construction of the relation between experimental fact and theory is usually considered in epistemological matters. Duhem seems to agree with this point; epistemological considerations, i.e. linking theory to experiment to speak about a theory's validity, are only relevant when the process of construction is complete. We shall return to this point later.

Quinean underdetermination

In I. Kant's philosophy of knowledge there is a separation of statements that describe matters of fact and those that describe matters of meaning independently from matters of fact. These are known respectively as the *synthetic* and the *analytic* statement respectively. Kant also separated the notion of *a priori* knowledge and *a posteriori* knowledge. *A priori* knowledge is knowledge that is or can be known before matters of fact are checked, while *a posteriori* knowledge is knowledge that can only come about as a result of checking the

[22] Duhem 1982, p. 206
[23] Idem

facts. Kant stresses that *a priori* does not mean 'available to us before a *certain* amount of facts is checked', but rather 'independent of *any* experience'.[24]

Kant's famous move was to introduce the *synthetic a priori*; knowledge about the world of experience which is nevertheless known before facts are checked. For instance Kant claimed that Euclidean geometry was *synthetic a priori*. Knowledge about space and time *presupposes* this geometry, and we can know that it must be the real nature of space to be Euclidean even before we actually make any inquiry into matters of fact.[25] With the creation of non-Euclidean geometries and the proof that these are relatively consistent with Euclidean geometry, as well as the eventual creation of physical theories in which non-Euclidean geometry was a description of space, thus countering the Kantian idea that Euclidean geometry is the true description of the nature of space, the whole idea of the *synthetic a priori* became suspect. Such strong intuitions as the Euclidean nature of space could be overthrown, so perhaps no facts about experience were *a priori*.

In his article from 1951, Quine dispensed with the other remaining distinction, that between the analytic and the synthetic. Quine is perhaps known best for his underdetermination thesis in the philosophy of science, but his abandonment of the analytic/synthetic distinction seems more fundamental and revolutionary. What is analyticity? What makes a statement analytic? Quine picks apart the most likely answers and concludes that for all its 'a priori reasonableness'[26] there is *no* distinction between synthetic and analytic. Analyticity can, for instance, not be a matter of definition, since definition itself hinges on the meaning of synonymous. So if we accept that analyticity is a form of synonymy, what does synonymous mean? Consider the phrases 'creature with a heart' and 'creature with kidneys'. It is clear that even though they may apply to the same class of objects –all creatures with hearts might have kidneys-, they do not agree in *meaning*, so we would not consider these phrases synonymous even if we could place an equality sign in between them.

Now consider two statements considered synonymous, 'bachelor' and 'unmarried man'. Quine says:

"There is no assurance here that the extensional agreement of 'bachelor' and 'unmarried man' rests on meaning rather than merely on accidental matters of fact…"[27]

And this is the main point of his rejection of the line between analytic and synthetic. What we consider synonymous is a result of the matters of fact concerning our understanding what the involved words mean, as well as what

[24] Kant, see introduction I, II, III, IV, V
[25] Kant, see SS3: Transcendental Exposition of the Conception of Space
[26] Quine 1998, p.290
[27] Quine 1998, p.288, 'extensional' means 'about the class of objects the phrase applies to'. So the extension of 'bachelor' is all men who are bachelors, which coincides with the extension of 'unmarried man'.

'synonymous' or 'equal in meaning' or 'analytic' mean. If we would happen to live in a world where 'bachelor' meant 'stone brick' there would be no synonymy, so the supposed analyticity of the two phrases *does* hinge on the matters of fact of our world.

This last remark gives rise to the idea of reduction, that is, the idea that all meaning can be linked or reduced to confirmation or disconfirmation by experience. The view that every meaningful statement can be translated into a statement about immediate experience is called reductionism, though this is a strong form of reductionism. This reductionism would 'save' synonymy in the sense that it is merely a matter of checking the facts regarding the experiences linked to the statements. This type of reductionism fails according to Quine because not every statement is translatable into direct experience. Take a simple sentence such as "A is equal to B", what does "is" mean in terms of direct experience? What would a word like "or" mean? It should be clear that not every statement is translatable in this way.

There is a weaker form of reductionism which holds that we can verify or invalidate statements to some degree of likelihood by means of checking a set of sensory events. For instance, seeing one black raven would increase the likelihood of the statement 'all ravens are black'. But this idea of verification requires some method of testing statements in isolation, and as Quine points out this is exactly what Duhem objected to. The conclusion is:

"[W]e noted […] a feeling that the truth of a statement is somehow analyzable into a linguistic component and a factual component. The factual component must, if we are empiricists, boil down to a range of confirmatory experiences. In the extreme case where the linguistic component is all that matters, a true statement is analytic. But I hope we are now impressed with how stubbornly the distinction between analytic and synthetic has resisted any straightforward drawing. I am impressed also […] with how baffling the problem has always been of arriving at any explicit theory of the empirical confirmation of a synthetic statement. My present suggestion is that it is nonsense, and the root of much nonsense, to speak of a linguistic component and a factual component in the truth of any individual statement. Taken collectively, science has its double dependence upon language and experience; but this duality is not significantly traceable into the statements of science taken one by one."[28]

As a result of this, we not only have the Duhemian freedom to choose which hypothesis to adjust in order to preserve correlation with experience of the entirety of our hypotheses, we also have the freedom to adjust the *language* we use in any part of this process. We can adjust logical laws, core words like "is" and the very definitions of the entities to be discussed in context of the realism

[28] Quine 1998, p.296

debate. This makes the problem of underdetermination even more forceful, for what our scientific hypotheses *mean* is underdetermined by experience; this Quinean underdetermination goes further than the underdetermination problem of whether the hypotheses in question are true or not when checking with evidence.

I think this puts Quinean underdetermination in a whole other class than what is usually discussed in the philosophy of science. In chapter one I introduced underdetermination and said that there are two related problems that are most prominently discussed as ‘the problem of underdetermination’; the problem of auxiliaries or holist underdetermination and empirical equivalence or contrastive underdetermination. It seems to me the Quinean underdetermination thesis goes further than both. Holist underdetermination states that, as no hypothesis can be tested in isolation, there is always room to adjust a group of hypotheses in such a way that they agree with previously disconfirming evidence. Quinean underdetermination *adds* that what the hypotheses *mean* is underdetermined by experience. This seems to me to be more fundamental. After all, we cannot talk about hypotheses and verification by experience without having some idea of what all the constituent terms mean. If we *do* have some way of knowing what is meant by all the terms, possibly we know this in an intuitive or unexplicated way, there is still room to adjust the meaning of the terms in the face of certain evidence. Finally, if we do have a fixed sense of meaning of the terms, there is still a holist underdetermination problem regarding the hypotheses in question. A similar argument could be made for contrastive underdetermination.

In Duhem’s physical holism there is a conventional aspect in the adjustment of the hypotheses in the face of disconfirming evidence. Now there is an additional conventional aspect in Quinean underdetermination in the adjustment of meaning attributed to the hypotheses. I think this can be illustrated with another geometry example. Consider a translation of a two-dimensional plane to the surface of a disc with a finite radius. Points on the plane remain points on the disc, and infinite unbounded lines on the plane become bounded curves with the endpoints on the edge of the disc. How can infinite unbounded lines become equivalent to bounded curves? Surely since one is infinite and the other finite there is no equivalence? Here comes the conventionality of definition; what exactly do we mean with ‘infinite’ and ‘finite’? We propose a definition of distance, as we are wont to do when faced with a disc which has a finite radius, the radius almost naturally proposing to us an idea of distance. Surely if we have a sense of distance we can define infinite and finite length in the sense of ‘amounts of unit length that fits in the radius’. It turns out the situation still has a conventional aspect. We are free to define the unit distance on the disc as going to zero as we near the edge from the point of view outside the disc, and if we use this definition of distance, then the curves on the disc will be of infinite length.

I hope the role of definition and meaning using this example is now clear. However, I can imagine this is not likely to strike any opponent of Quinean underdetermination as a very impressive example, given that it is about artificially constructed mathematical concepts. I wish to make clear this example is easily adapted to our reality, and thus made more cogent. Consider how we are going to determine whether or not we live in an infinite unbounded three-dimensional universe or on a bounded spherical universe.[29] The measuring devices we use to determine such questions also use specific definitions of distance. How are we to determine what the 'real' measure of distance is in our universe? It could very well be we live in a bounded spherical universe, if matter within the sphere becomes more and more dense and compressed in the outward direction as we reach the edge. To our experience, the universe would be unbounded and infinite. What is the true nature of our universe?

I believe a similar issue was at stake in the case of Lorenz versus Einstein, where the question was about a definition of 'rest' and whether or not the Lorenz contraction was an inherent effect of the nature of space-time or a result of molecular forces acting upon a body not at rest. There is in special relativity also the case of the definition of simultaneity. Simultaneity is defined using a roundtrip of a light beam, but it can be considered an issue of convention whether or not light has the same velocity going one way in the roundtrip as the other. Due to principles of symmetry or parsimony the general tendency is to consider the velocity the same going one way or the other. But actually what is the case seems to be underdetermined. Appealing to principles of symmetry or parsimony to escape the underdetermination problem should require spelling out of such principles and providing justification for them. This will be the subject of part 2 of this thesis.

So we see that Quinean underdetermination is in principle a real problem for physical science. That leaves open the question if, like in the Duhemian case, there is room for an extralogical process which rationally justifies choosing one hypothesis over the other in the light of Quinean underdetermination. Quine himself considered the extralogical process to be purely one of pragmatism and rejects realism, while Duhem appeals to *bon sense*. In part 2 of this thesis we will consider the value of such extralogical processes but first we will put the discussion in a more modern light by taking a closer look at H. Collins' book 'Changing Order'.

[29] For simplicity let's ignore time as a dimension.

Collinsean underdetermination

In the book 'Changing Order' H. Collins gives his appraisal of replication and induction in scientific practice. I think a large part of his analysis can be understood in terms of underdetermination. Collins himself does not use the term underdetermination, but this is not surprising since he does not have a background as a philosopher of science. Collins' book is one of the seminal works in what is called 'the Sociology of Scientific Knowledge' [SSK]. In the field of the philosophy of science SSK is infamous for its relativism. That is, the tendency of academics in this field to claim that there is no *rational* solution to the problem of theory choice, that there is no *objective* scientific method and that any decisions made by scientists and any worldviews they may have are merely a result of sociological or cultural processes. To this last remark I must add that like in all philosophical positions there is a varying degree to which authors considered part of SSK believe sociology or culture is involved. For instance, if we were to include Kuhn in SSK, we would find that despite his claim that theory choice 'cannot be resolved by proof', there is nevertheless a rational choice possible due to the collective judgment of scientists: "What better criterion could there be [...] than the decision of the scientific group?"[30]

While Collins is obviously within the tradition of SSK, I think his book merits a fair treatment as he raises some points which are significant to the discussion on underdetermination. In fact, a few of the points in his book are so similar to a form of Duhemian underdetermination I think these points can be described as a modern version of Duhemian underdetermination. The most striking claim is that there is in the practice of science a problem called the 'experimenter's regress', and I think this problem can be understood as a form of underdetermination in the reality of scientific practice.

The experimenter's regress exists ultimately, so says Collins, because of the problematic nature of inductive inference as explained in chapter one. In this chapter Collins uses a few entertaining examples to explain the problem of inductive inference, and one of these is worthy of review:

"What this involves can be explored by looking at a much more straightforward sequence – the numbers '2,4,6,8'. Imagine being asked to continue this sequence in the same way. The immediate answer that springs to mind is '10,12,14,16' and, to all intents and purposes, this is indeed the 'correct' answer.

But how do we know it is the correct answer? It cannot be simply a matter of following the rule 'go on in the same way' because this rule allows for a number of possibilities. For example it allows '2,4,6,8,10,2,4,6,8,10,12,2,4,6,8,10,12,14' [...]

[30] Kuhn 1998, p. 102

the instruction could [...] also allow for 'who do we appreciate?' as the continuation."[31]

This problem is then subjected to sociological analysis. The solution that follows from this analysis is that agreement comes about through 'forms of life' within social groups. These groups have social conventions, and it is these conventions which can solve the problem. For instance, by social convention it is understood that the proper answer to '2,4,6,8,... what follows?' is '10,12,14,...' or in other contexts 'who do we appreciate'.

Some philosophers do not find this solution satisfactory and remark that Collins dismisses modern 'rational' alternatives all too readily or that he does not even consider them, criticism that is similar to realist arguments against underdetermination. For instance, John Norton explicitly challenges relativism on the grounds that it 'neglects the literature in induction and confirmation'[32] and this seems to be the prime concern in Mary Hesse's review:

*"Science does not depend on algorithmic rules of methodology, nor on necessary and sufficient conditions for what counts as replicability, and, in general, science need not be supposed to exhibit one-to-one correspondence with objects and regularities in the world independently of human categories and classifications. But the need to modify the view of science in this way has **long been accepted**, and **there are many revised accounts** that do not entail radical conventionalism."*[33]

While this may be a legitimate concern, I think Collins' response is already contained within his 'awkward student' experiment: Imagine two people fulfilling a different role; that of the Awkward Student and the Instructor. The Instructor has the task of giving some set of instructions so that we continue '2,4,6,8' in the proper way. The Awkward Student must try to interpret these instructions so that he follows those using logical rules of application, but does not come up with the intended result. For instance, given a rule of 'continue in the same way' the Awkward Student could interpret 'in the same way' as repeating the sequence itself so that he gets '2,4,6,8,2,4,6,8…'. The Instructor gets to respond every time to adapt the instructions to what the Student does. Eventually, either the Instructor gives up, or he inserts a rule that trivially solves the problem. In this case such a trivial solution could be "continue with '10,12,14…', which would defeat the purpose of the puzzle. However, even this solution is open to Awkward Student misinterpretation since unless a rule is specified, he could misunderstand 'with' and 'continue'.

This last remark will sound similar to points Quine raised, but there is an important difference. The example is about a *human concept*, namely a number

31 Collins 1992, p.13
32 Norton 2008, p. 26
33 Hesse 1986, p. 725

puzzle with an intended solution which exists by the grace of our being, so this is not obviously equivalent to a situation in which we compare sentences to an objective reality outside our own minds. That is to say, it *might be* equivalent in a Quinean way but that would require additional support. The important thing to note here is that Collins shows how Quinean underdetermination is a very real part of life, at least when dealing with human concepts, and that he proposes that we all have a capacity to solve the problem but not in a logical way. This isn't surprising as far as I am concerned; it is merely a matter of solving a conventional issue with a convention. A convention is involved when posing a puzzle like '2, 4, 6, 8, continue in the same way', for the meaning of the words and the meaning of this entire puzzle, is a convention. I will return to this point in my criticism of Quine.

Let's return to the criticism of philosophers. If we imagine the Awkward Student trying to apply the rules and methodologies given by philosophers of science then we can understand why Collins does not even consider the 'rational alternatives'. Collins considers these rational attempts at a solution as *rules* in one form or another and thus these attempts are *a priori* thwarted by his Awkward Student, since the student will always find a loophole. However, I'm not sure it is fair to say that philosophers propose their solutions to problems of inductive inference as rules and methodologies. I will be reviewing such criticism in chapter 2. It is outside the intended scope of this chapter. For one, it concerns the question whether or not a proposed loophole is *rational* to accept.

The game of awkward student seems to add a new perspective on inductive inference and it leans heavily on Goodman's Gruesome New Riddle of Induction in Collins' account. The bottom line of Goodman's Riddle is that regularity can be seen *anywhere*. For instance, an emerald is understood to be *grue* if it is green before time X and blue thereafter, where X is some time in the future. The point being that we can see both the regularity of all grass seen before X as greenness or grueness. How do we know an emerald is actually green and not grue if we can perceive both regularities? While this grue concept might seem artificial, Goodman's point is that we have no *logically* compelling reason whatsoever to pick green over grue, similar to how Hume pointed out that there is no logically compelling reason to presume a regularities' occurrence a next time, and that we are logically free to call an emerald grue, since given past experience whether an emerald is grue or green is equally supported. Instead of 'seeing' the greenness of emeralds as regularity, we might as well have seen the grueness of emeralds as a regularity.[34] Collins' solution to Goodman's Riddle is that the community *decides* to pick green over grue; this is a social convention.

[34] Vickers 2011

While Collins states that this is a problem of inductive inference I think it should not be considered a problem of inductive inference but one of underdetermination. As mentioned before I make a distinction between induction and underdetermination and claim that these types of inference have their separate though related problems. I think there is some confusion of what is generally meant with 'induction' and 'underdetermination' mostly due to unclear use of both terms, and I think that Collins falls prey to that error here. It is important to note that philosophers of science have tried to do exactly what Collins claims is impossible; they have tried to come up with inductive rules of inference that solve our philosophical problems and once and for all provide a steady base for sound inference from evidence to theory.

The Awkward Student can misinterpret any given rule *not* because he does not have access to the *proper* inductive inference method, but because given the available evidence *the solution to the puzzle is underdetermined*. It remains a question whether or not inductive inferences allow us to avoid this underdetermination. Philosophers of science propose different kinds of inductive inference methods and present them as *the solution* to the kind of problem Collins describes. If they *would* have an inductive method that would be accepted as or demonstratively or rationally sound, and not *just* accepted because of social processes, then the problem Collins points out would not arise at all. But I must add that induction is not the only possible answer; in chapter 2 I will look at *adduction*.[35]

In chapter two Collins argues that because of the problem discussed above the replication of experiments is problematic. The problem is how to infer similarity or difference between experiments. Scientist A doing an experiment here and scientist B trying to repeat the experiment there are, intuitively speaking, performing more similar experiments than the same scientist A and a gypsy C reading goat entrails to get the results the scientists A and B are after. This seems intuitively clear, but how can we properly formulate this difference in the sense of philosophical rules? As a corollary: how do we determine which elements of an experiment matter when determining this similarity? Collins invokes a metaphor in which philosopher mice try to formulate the rules which guide the steps involved to come to such a rule of similarity, but in his analysis the mice have serious problems at every step. Collins' solution is unsurprisingly a sociological solution.

I am however unconvinced as how to connect this with the problems discussed above. It seems acceptable to me to just take the analysis as is, and I do not see the logical connection between the grueness problem and similarity in experimentation. Here then is the analysis as is:

[35] Both terms will be explained below.

"For an experiment to be a test of a previous result it must be neither exactly the same nor too different. Take a pair of experiments – one that gives rise to a new result and a subsequent test. If the second experiment is too like the first then it will not add any confirmatory information. The extreme case where every aspect of the second experiment is literally identical to the first is not even a separate experiment."

"Confirmatory power, then, seems to increase as the difference between a confirming experiment and the initial experiment increases."

"If we move back in stages toward lesser degrees of difference we can now see that the situation steadily improves. […] Thus, if the gypsy had used some old technical equipment rather than goat entrails [to obtain the same result as an experiment], it would look a little better. If the gypsy is replaced by a high school student it looks still better. […] If the high school student had used a good apparatus, then things would be better, and likewise if it were a first-rate physicist who had used the poor apparatus."

"This is because if a second experimenter fails to see a claimed result, differences of design between first and second may be invoked as the cause of the failure; it will be said that the second experiment has not been done according to the instructions. Thus the strength of a disconfirmation goes up as the second experiment approaches identity with the first."[36]

Confirmation increases as confirmation is found in increasingly different circumstances, but will decrease if found in questionable circumstances. Disconfirmation decreases if a test is too much unlike the tested. This seems to me a sound analysis by itself. The question is: at what point does a potential confirmation or disconfirmation turn into an actual confirmation or disconfirmation.

In the following chapter Collins illustrates the problem with a case study in which he followed experimenters trying to build so-called TEA-lasers. He describes which steps they took and how they decided which aspects of the laser were important for the success of the experiment. Like in the following case studies it turns out even within what is considered to be one and the same experiment yet executed by different experimenters, controversy arises through difficulty in controlling the instruments and uncertainty about the validity of the results. We see in practice the difficulties the analysis above points to. In the case of the Jumbo version of the TEA laser, a 'copy' was built but failed to perform as the original. As the experimenters investigate the cause of their problems, they find differences between the original and the copy that were not perceived before, as well as literal differences that are not perceived as differences, because they are deemed irrelevant.

[36] Collins 1992, p. 34-36

The problems are in practice settled in due time. In the meantime experimenters seem to gain an intuitive understanding of their instruments, and Collins likens this way of dealing with practical problems in experimentation to learning a craft. Experimenters learn to experiment just like blacksmiths learn to smith iron – a point I. Hacking will appreciate for sure[37] -. There is a great deal of what he calls 'tacit knowledge' involved. Collins takes the concept of tacit knowledge from Michael Polanyi; it means '[the] ability to perform skills without being able to articulate how we do them'.[38] An example is learning how to ride a bike; even if we don't know how exactly riding a bike works, we learn to do it without having access to the governing principles –if there are any-.

According to Collins, the reason experimenters don't immediately get a clear cut view of their experiment and the reason why controversy like mentioned above is possible in the first place is the experimenter's regress as mentioned earlier. Experimenters try to build instruments to do experiments on some kind of phenomenon for which they are perhaps searching confirmation; the phenomenon in question is not directly observable. The regress is a direct result of the fact that the workings of the phenomenon and the inner workings of the instrument are *a priori* unknown. The experimenter must have an idea of what the correct outcome of his experiment is or of what the results of the experiment mean in terms of the theoretical framework involved, and to do this he must have an idea of what the instruments do and what constitutes success in terms of observable effects of the instrument. But what the correct outcome is, and what the correct observable effects are, depends on whether or not the phenomenon is actually real and how it interacts with the instrument. And we want to find out whether or not it is real and how it interacts with the instrument *through* the experiment. So there is a degree of circularity involved if we are just considering observable effects of an instrument.

This seems to be very similar to Duhemian holism. The experiment depends in a crucial way on our *a priori* conceptions of the parts of reality involved, so a clear cut outcome, and correspondingly a theoretical meaning of the outcome of an experiment, is impossible. It seems intuitively plausible that this *would* be possible if an experiment is being replicated, since the supposed outcome and intended theoretical meaning are known from experiments beforehand. For instance, think of experiments meant to gauge instruments. Even if we have access to such experiments this does not avoid the problem; if the assumptions of the first experiment are wrong, then replications similarly rely on false assumptions even if the intended outcome can be reproduced. In other words, if we rely on other experiments this does not avoid the experimenter's regress; the regress problem just shifts to the other experiments.

[37] This is similar to Hacking's general view of science, see for example Hacking 1985.
[38] Collins 1992, p. 56

For example, we try to measure the Higgs boson. Assume that our theories tell us that it is supposed to give reading P on instrument Q. The regress problem is that we want to find out if the Higgs boson exists but at the same time it is a question whether or not the Higgs boson actually gives reading P on instrument Q, because we still need experimental confirmation of this theoretical statement. Even if we have countless experiments that show us that certain other particles give reading P and if other particles have the same property (property R) as the Higgs boson, R being assumed to give reading P, this only shifts the problem. Firstly, in the particular case with the Higgs boson the circumstances could be so different as to not give reading P even if the Higgs boson does have property R. Secondly, the experimenter's regress is now applied to the experiments in which other particles give reading P, and the experiments that are supposed to show that the other particles have property R. Thirdly, the experimenter's regress is now applied to the experiment that is supposed to determine that property R gives reading P. The problem just got worse!

In practice, it turns out that in the case of replication, as Collins shows, the outcome of the experiment to be replicated is unclear because new experimenters might call into question the understanding of the experiment, or because the new experimenters lack certain tacit knowledge and fail to reproduce earlier results. We cannot isolate the theoretical workings of the instrument from the outcome of the instrument, thus we have to deal with every theoretical statement involved at once and can only make claims about the group as one. This is similar to what Duhem said, but Collins now also explicitly denies the possibility of verification of an isolated hypothesis.

To come back to my earlier point of criticism, I doubt that the problematic nature of inductive inference should specifically play a role here. I think that Duhemian holism follows from the nature of science, theory and experiment, and that no account of problematic inductive inference is necessary to justify Duhemian holism. Duhem certainly doesn't base his holism on the limited nature of inductive inference. There *could* be a rule of inference that constrains the inference from observation to theory in such a way that isolated evaluation of involved hypotheses is possible, but I am not sure for reasons to be explained below that such rules are properly described as a *inductive* forms of inference.

So experimenters must judge the status of the instruments involved, but judgment *cannot come from the experiment itself*. Collins states that the solution to the regress must come from the outside. The case studies in his book, apart from the TEA laser there are two other examples, illustrate the particular ways in which the regress, as Collins sees it, have been resolved. Collins spends the sixth chapter generalizing his findings, and his response to the problem is as mentioned above a sociological solution. In this chapter Collins tries to propose a new view of science, and while he has some interesting insights I think presenting this as *the* solution to the problematic nature of inductive inference goes a bit too far, not to mention my concern that induction is not that important here. If inductive inference is as problematic as Collins presents it then surely the step from finding the locus of the problem in the relation between theory and experiment to concluding that sociological processes are the end-all solution should be highly suspect since it in turn requires a kind of 'inductive' inference. To put it in terms of *underdetermination*: given the evidence -namely problems with the relation between theory and experiment- what the proper solution is -in particular whether or not this is sociological solution- remains underdetermined.

Collins' analysis of experiment and inductive inference is solid and I think he has a point when he concludes that scientific knowledge does not come about through use of certain algorithmic rules. He would certainly not be the first philosopher to point this out, as Feyerabend already denied any universal methodology 10 years earlier in the book *Against Method*. But it does not automatically lead to conclusions like 'it is not the regularity of the world that imposes itself on our senses but the regularity of our institutionalized beliefs that imposes itself on the world.'[39] He could have, given the evidence, followed Feyerabend and concluded that there is no single objectively superior set of algorithmic rules, but that there are many sets of rules, each used –and superior- in different cases whenever the need arises and that there are even methods of progress which do not take the form of rules at all.[40] Collins could also have followed Norton who claims that scientific knowledge is, metaphorically, like a stone arc, each 'stone' of it resting crucially upon others, like each stone of an arc can only stay up if it is wedged between its neighbours, but together the entire thing stays up[41]. Not to say that Collins' conclusion is *demonstrably invalid*, but like in experimentation the conclusion *is underdetermined* by the available evidence, and to state that a sociological solution is *the* solution with such certainty requires a leap of faith at this point.

To use an analogy with Goodman's riddle, it is logically possible for green or grue to be a concept describing the real colour state of an event, and it is possible that both describe the actual state of the colour of emeralds, like it is possible that both the sociological and traditionalist philosophical solutions are the proper solution to our problems of 'inductive' inference. It is, however, not obvious *which* of green and grue are the only true description of events, just like it is not obvious whether or not the sociological processes described by Collins

[39] Collins 1992, p. 148

[40] Feyerabend does not state this last part explicitly in Against Method; he merely rejects that there is a *single* set of rules which objectively govern scientific progress.See Feyerabend 1993, particularly chapter 17 and 19.

[41] Norton 2009

are the real benefactor to solving our problems, or that it is a good old-fashioned use of Bayesian algorithms or Carnap logic or other self-styled *rational* alternatives. What is obvious, as far as I'm concerned, that not both green and grue can be true of one and the same thing -except for multicoloured objects and other trivial exceptions. However, the analogy breaks down at this point; it might be the case that many proposed solutions to the problem of induction work together to shape our convictions, such as sociological processes being guided by Bayesian processes, while in the gruesome argument either 'green' or 'grue' is true.

Taking underdetermination seriously

Now that I have reviewed various representations of underdetermination and the original arguments by Duhem and Quine I can propose a proper representation of underdetermination. Recall the two forms of underdetermination as occurring in the modern discussion. Holist and contrastive underdetermination both seem to raise the immediate question of their validity. Why should there always be auxiliaries to adjust so that we can reconcile any evidence with any theory? Why should their always be possible alternatives in the form of empirical equivalents? This is exactly the points Laudan, Lipton *et alii* raise.

Perhaps this is just a strong reaction to Quine who stated so boldly that '[a]ny statement can be held true come what may, if we make drastic enough adjustments elsewhere in the system [of beliefs]'[42] but it seems odd that so many responses to underdetermination have focused on holist or contrastive underdetermination. These responses seem to boil down to simple scepticism about their validity, but do not go into the origins of the argument. These arguments *do* find further justification in both Quine and Duhe. Duhem did not merely claim holism or the possibility of empirical equivalents, *he gives reasons for these claims*, and the reasons are *thoroughly compelling*. There is a lot in this discussion that does not take into account Duhem's original analysis. Similarly for Quine, as he based the rejection of one of the two 'Dogma's of Empiricism' partly on Duhem's account. We find no rejection of these arguments in the modern discussion on underdetermination, nor reasons why we should think the analysis by Duhem or Quine is invalid.[43]

For this reason we will focus our attention on the original accounts. So we have one strong –so far- version of underdetermination:

Duhemian Underdetermination (DUD): *There can never be absolute falsification of a single theoretical statement, since testing such a statement, via experiment, requires a network of other theoretical statements. These additional*

[42] Quine 1998, p. 296-297
[43] See for example Laudan and Leplin 1991, Laudan 1998, Norton 2008 and Kukla 1996 and 1994. All of these are aimed at contrastive or holist underdetermination.

statements can for instance come from other theories, which are needed to indirectly link the tested statement to observable reality, or they can come from theory needed to use or understand an instrument. The result is that a theoretical statement cannot be rejected in isolation; only a group of theoretical statements as a whole can be rejected.

As a result of DUD, it is not immediately clear how general underdetermination should follow, since DUD itself still allows verification of hypotheses. Thus a *constructive* view of science is still possible, though the question is raised how to get rid of falsehoods since falsification is problematic. In particular, it isn't clear how exactly from DUD holist and contrastive underdetermination should follow, though obvious intuitions exist. We enhance DUD with another form of underdetermination as found above:

Collinsean Underdetermination (CUD): *There can never be absolute verification of a theoretical statement, since testing such a statement invokes the statement in question or statements that it fundamentally relies on due to the indirect nature of testing such statements. The method of verifying a theoretical statement requires an idea of what a successful experiment would be, that idea itself is dependent on the truth of the theoretical statement. A way out can only come from outside; relying on other experiments does not work because they themselves are vulnerable to CUD.*

Note that CUD is not necessarily a new thing. It can be derived from the exact same analysis of the relation between experiment and theory as Duhem has given. That Duhem himself does not explicitly do so may seem odd in retrospect, but we shall accept CUD as filling this gap.

Taking these statements together there is, due to the indirect nature of testing theoretical statements, no way to find absolute falsification or verification of a theoretical statement. We combine the two forms of underdetermination to find the thesis of underdetermination which we should take seriously in the realism debate. This form is the new 'lord' of underdetermination theses. Thus, the new form of underdetermination looks like this:

DONUD ≔ DUD + CUD

DONUD recognizes the problematic nature of the relation between theory and experiment. It admits that instruments provide an indirect means of investigation. As a result of this, there is logically a 'web' of theoretical statements which are *loosely* connected with purely empirical statements. The theoretical statements *cannot be reduced* to purely empirical statements, and this in turn invalidates the logical positivist framework of science. Through experiment there is, however, a set of empirical statements connected with certain theoretical statements, though not necessarily in a 1-to-1 or exhaustive relationship. The observable-unobservable dichotomy plays a vital role here.

DONUD creates the following image of experimentation: $E + I \leftrightarrow T + C$, where E is the empirical outcome of the experiment, I is the interpretation of the empirical outcome, T is the theory governing the experiment and C various other conditions particular to the experiment. The $\leftrightarrow$ sign indicates that E + I and T + C should provide the same thing, that is, the interpreted outcome of an experiment should be the same as the theory's prediction with given circumstances. E should be seen as expressing the empirical outcome in purely observational terms. E does not say 'there is an electron at X' but something like 'voltmeter P's indicator points at 4V'. This is in turn interpreted via our current theories to say something like 'there is an electron at X', which coincides or not with what our theory says happens under conditions C.

If an experiment fails or succeeds, this does not mean T is false or true respectively. I, T and C may be different from our expectations without our knowledge, creating error or deluding us into thinking the experiment was a success. The point of anti-realism based on DONUD is that we simply cannot know, and thus should have an a priori sceptical attitude towards all theoretical statements.

As an example, imagine an experiment in which some chemical compound is investigated; the goal is to determine its molecular constituents. The investigators use a device called a *mass spectrometer*. The outcome of the experiment is that the compound consists of this much element X, this much element Y and this much element Z. We can now look at the statements involved in the way above. The investigators use terms like 'molecular constituents' and 'element X Y and Z'. The truth of such statements is underdetermined according to DONUD, but we can still look at the *empirical* results. There are *observable facts* involved, namely a graph plot with spikes at certain points. We connect theoretical statements to such facts; we interpret them to imply something about a world *out there* yet not directly accessible. In this case a spike at a point labelled X with a certain height is interpreted as 'this much of element X'. But we must not confuse these interpretations with empirical facts, for *they are not facts* according to DONUD.

In addition, I think we can reject Quinean underdetermination (QUD). QUD relies on our ability to adjust the meaning of terms, perhaps even on a whim. But the meaning of a term for observable facts is just a description; it is a name given to a certain experience. For instance, we have given a certain colour sensation the name 'green'. Why would we, when we have given something a name and agreed upon it amongst ourselves, change the meaning of such a term? Quine uses the example that 'bachelors are unmarried men' is not analytic, because the terms 'bachelor' and 'unmarried man' might have meant something different, for instance in their extension. But this strikes me as a trick. We understand ‘bachelor’ and ‘unmarried man’ to be the same precisely because both are the name *we have given* something which takes place in our experience. Quine is right when he says that this means such statements being synonymous is a result of matters of fact, and that ‘bachelors are unmarried men’ is a synthetic statement in that regard, but that does not mean we should have the freedom to rearrange any language terms as we see fit. After the facts of the matter, for instance ‘bachelors’ happening to mean ‘unmarried man’, are established we should have a relatively stable use of language. Quine only shows that language is partly conventional, not that once the conventions are set there is fundamental underdetermination due to language.

QUD does still apply for theoretical statements, but as long as DONUD is valid for the theoretical level QUD is an afterthought. For instance, what ‘electron’ means could change, and it has through the years. But this should not be strange to those who accept DONUD, for the whole existence of a thing like an electron is underdetermined, let alone its properties. If DONUD is valid, it should not be a surprise that in the history of science theoretical terms have come to mean different things. DONUD provides the room to allow this sort of change in the meaning of theoretical statements.

Underdetermination is not a problem of induction

Underdetermination should be seen as a different problem than ‘the problem of induction’ as understood in terms of enumerative induction. Recall that enumerative induction is about inferring from particular instances of a statement to a future instance of that statement. If we allow enumerative induction, that is we accept it as a valid method of inference, then logically we can infer from particular instances of a statement that the statement will hold for all future instances. So for instance if we see multiple times that an A is B then we can conclude inductively that all A’s are B, I’ll write (all observed A’s are B) $\rightarrow_{ind}$ (all A’s are B).

The problem of underdetermination in the philosophy of science is another matter. In science we do not usually try to infer from particular statements to its generalized statement. We instead try to infer from particular statements to generalized statements from which the particulars, sometimes even other particulars, can be deduced in turn. For instance, we do not infer from repeated experiments in which two electrically charged particles of a certain type placed together start moving in such and such a way that *all* particles of that type always move in such and such a way if placed together, but we instead infer that there is a *law of Coulomb* which governs all electrically charged particles, and that there is something like a ‘charge’ and a ‘force’ involved. This law is a different beast entirely, because it governs many more phenomena.

The types of inference where underdetermination plays a significant role involve more than just inferring (all observed A's are B) $\rightarrow_{ind}$ (all A's are B); it involves inferring (all observed A's are B) $\rightarrow_{ind}$ (principle C from which it can be deduced that all A's are B). Usually principle C can be used to deduce many more statements than (all A's are B). According to DONUD, this last form of inference allows that there are multiple, perhaps inconsistent, principles C' possible that could all deductively entail (all A's are B). It is the point of this thesis that this is a problem particular to inferences regarding unobservable events. If it were about observable events, we could solve the underdetermination problem by observing, leaving only one observed C.[44] This turns underdetermination into a *practical problem*, that is, a problem we face every day and we can solve in practice by applying trial and error. But it is not a *principal problem* like it is for unobservables, that is a problem which is fundamental for the appraisal of truth. If we look at what Duhem and Collins and even what critics like Laudan have written about underdetermination, and we take them to be good examples of underdetermination, we find that underdetermination is well represented as being something different from enumerative induction. In fact, if we grant that multiple principles C deductively lead to (all observed A's are B) we see that underdetermination is a problem even if we allow, contra Hume, inference in the form of enumerative induction.

In practice, I think whether or not enumerative induction is problematic is not an important problem. We use that form of induction on a regular basis and it seems to pay off. Empirically speaking it seems that in certain cases enumerative induction is a reliable way of inference. The more interesting question is about *when* enumerative induction is reliable and *when it is not*. Again, empirically speaking sometimes it seems to work well, other times it fails miserably. For instance it seems clear that when talking about me jumping out of a window we can safely infer that I would fall to my death, yet when we are talking about me showing up in the students' room to write on my thesis there is no basis for inference based on enumerative induction, as I show up almost as often as I don't in no particular pattern. What makes these situations so different? How can a rule of inference separate the cases?

An answer to such questions might well be able to solve the underdetermination problem. Philosophers of science of the last century have tried to provide such answers. What seems clear is that purely deductively speaking there is little room to maneuver, so philosophers of science have sought inductive or other ways of inference to resolve the problem. If, like I mentioned when I discussed Collins, there is a way for us to accept a certain inductive inference, we may be able to use that form of inductive inference to

[44] More justification for this position is found in part II.

justify a form of realism. Other routes of inference will be investigated in part II of this thesis.

As a preliminary remark, I will warn that there could be a bit of confusion regarding the status of ‘inductive’ inference. To avoid confusion I will separate different kinds of inference. I will refer to the kind of inference in which underdetermination plays a role, as discussed above, as ‘adduction’. In the discussion above any inference rule that allows us to choose one of the *principles* C would be an *adductive* rule of inference instead of an *inductive* rule. There are rules of inference which are properly called inductive, of which the most popular example is Bayesianism. I will discuss Bayesianism and its capability to deal with underdetermination in part II.

Another characterization of DONUD

P.K. Stanford gives a modern characterization of underdetermination which he claims is a credible form of underdetermination. I will now discuss his proposal and compare it with DONUD, and we will see that he has in mind a similar type of underdetermination, not based on the tired old holist or contrastive underdetermination. Stanford rejects arguments from anti-realists and realists alike.

Stanford means to show that classic defence attempts of contrastive underdetermination in which we can always create logical constructions to act as possible rivals for a scientific theory ‘amount to no more than a salient presentation of the possibility of radical or Cartesian skepticism.’[45] Stanford however argues that this is beside the point. On his reading, undetermination was originally thought to be a separate problem from general issues of skepticism. I agree with this assessment as should become clear in this thesis. Stanford points to an interesting form of underdetermination; he calls it ‘Recurrent, Transient Underdetermination’[46], which entails that ‘there might simply be garden-variety alternative hypotheses, *not yet even imagined or entertained by us*, but nonetheless consistent with or even equally well-confirmed by all of the *actual* evidence *we happen to have in hand.*’[47] I think such a form of underdetermination is a natural consequence of DONUD and that is precisely formulates why DONUD is a real problem. Note that as Stanford points out, Duhem has said this:

“[L]et us admit that the facts, in condemning one of the two systems, condemn once and for all the single doubtful assumption it contains. Does it follow that we can find in the “crucial experiment” an irrefutable procedure for transforming one of the two hypotheses before us into a demonstrated truth? Between two

[45] Stanford 2001, p. S3
[46] I will refer to this as RTUD.
[47] Stanford 2001, p. S7

contradictory theorems of geometry there is no room for a third judgment; if one is false, the other is necessarily true. Do two hypotheses in physics ever constitute such a strict dilemma? 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?"[48]

In other words, I follow Stanford to dismiss radical scepticism in the sense that we give no reason why it should be taken seriously. Instead, we should apply a limited form of scepticism, in which we do not regard merely any logically trivial construct rival but only serious scientific possibilities. This is in fact enough reason to adopt underdetermination, so we can adopt a form of underdetermination which avoids radical scepticism. Logically, a defence of radical scepticism is still possible, but I will attempt no such course of action here.

Stanford wonders if there are actual examples of situations in which this predicament turned out to be true. The idea is that if RTUD is just a theoretical possibility which does not bear out in actual practice, it is not convincing despite being theoretically sound. This is because if cases of RTUD cannot be found in actual history there is little support for the idea that 'there might simply be garden-variety alternative hypotheses, *not yet even imagined or entertained by us*, but nonetheless consistent with or even equally well-confirmed by all of the *actual* evidence *we happen to have in hand*.' Stanford provides enough evidence himself to conclude that situations of RTUD have occurred frequently enough to warrant a *New Induction over the History of Science*; the historic record suggests that RTUD has always been a problem, so we expect it to be for our current theories. I have little to add to this except that I think accounts of actual practice of science, like that of Collins, show that RTUD is even a problem on the micro-scale of work in the laboratory. For example, in the case of the Jumbo laser the solution to make the laser copy working was something the experimenter himself didn't even think of. [49] Or we can think of the history of models for the atom. It is a simple fact that these models changed over time, yet earlier models were used with experimental success. Thus, it is a simple truth that plausible rival theories are in practice found, with the proviso that we need look beyond the short term.

Finally, Stanford writes the following:

"The New Induction will nonetheless disappoint a great many champions of underdetermination, for the historical record offers at best fallible evidence that we occupy a significant underdetermination predicament, rather than the sort of proof

[48] Duhem 1982, p. 189-190
[49] Collins 1992, p. 62

that advocates have traditionally sought (and I have been unable to do more here than suggest that this is indeed the verdict of the historical record).”[50]

I think that if Stanford would have accepted Duhem’s analysis as evidence beyond the New Induction, he would have concluded that Duhem’s analysis is in fact enough proof when presented together with the New Induction argument. The proof that Stanford seems to require is given in DONUD. So in my view the combination of Duhem’s analysis, Collins’ experimenters regress and Stanford’s New Induction is enough evidence that DONUD as presented is a serious problem, not only because of the theoretical review of the relation between theory and experiment, but also because it agrees with the historical record, and can explain some peculiar details about the practice of science as explained in Collins’ case studies.

Conclusion: What is underdetermination?

Concluding this part of the thesis, we find that if we are to take underdetermination seriously, we should look at DONUD, and not focus solely on contrastive or holist underdetermination. DONUD says something about our ability in practice to gain knowledge about the unobservable in addition to our ability to gain this knowledge a priori. That is, DONUD is not only a fundamental problem for realism because it is a serious obstacle for gaining knowledge about the unobservable, it is also a practical problem in the sense that it is the reason why experimentation is a road of trial and error, with emphasis on the error. In that regard DONUD is stronger than both contrastive or holist underdetermination, because those two forms of underdetermination entail that there is a general availability of either empirical equivalents or adjustable auxiliaries. It should not be surprising that scepticism about both forms of underdetermination often appears in the form of a rejection of their *a priori* credibility. In contrast, Duhem and Collins explicitly point out why there is respectively no falsification or verification of an isolated hypothesis in experiment. These analyses are sound, even if one can disagree, as I have, with how Collins embeds this in a framework of social decision making. Instrument and theory are entangled in such a way that underdetermination constitutes a problem in the practice of science as well as a problem of justification of the truth of hypotheses. This is reflected in historical case studies such as Collins’, and every scientist who has ever done experimental work should find DONUD familiar; it is the reason why experiments usually don’t work out the way they are imagined even if it is ‘merely’ a case of replication.[51] Struggling to find the

[50] Stanford 2001, p.S11

[51] The Jumbo case illustrates this well in my view.

intended solution of an experiment is part of the process of solving the problem of underdetermination in practice.

This leaves the question whether or not DONUD is a serious problem for realism or turns out to be not so malicious after we abandon the need for deductive proof. In other words, are there *reasonable* ways out of DONUD? Is there an inductive or adductive rule of inference which allows us to escape the problem in a reasonable way? In part II I will address such questions, as well as criticism about the underdetermination thesis. We will see that DONUD will have to be adjusted slightly in order to hold ground against some of the arguments put forward by realists.

PART II: CAN DONUD BE SOLVED?

The modern discussion on underdetermination

With a clear picture of what kind of underdetermination I take seriously I can focus on the arguments for and against it and in particular I can evaluate the relation between underdetermination and anti-realism. At the end of this part of my thesis I hope to have a clear view of underdetermination as well as of a form of anti-realism soundly based upon said underdetermination. In order to do that we need to shake ourselves free from the intellectual baggage which 100 years of the realism debate has put upon realism and underdetermination. I will dismiss some common forms of anti-realism and I will have a positive attitude towards arguments against these common anti-realisms. These arguments will help to provide a stronger view of anti-realism than some of the currently available alternatives. In a different perspective this could also be seen as a sharpening of current versions of anti-realism, most notably constructive empiricism.

I will pay attention to different kinds of inference as mentioned at the end of part I. Sometimes it is claimed that accepting the 'underdetermination thesis' automatically leads to radical skepticism regarding all kinds of -inductive- inference.[52] It is proposed that if we allow underdetermination as a serious problem we can have no meaningful knowledge whatsoever, for example about whether or not we'll fall to our deaths if we jump off a high building. Sometimes such arguments seem to suggest that we might as well try if there is no certainty due to a Humean kind of underdetermination –that the future is underdetermined by the present and past. We can dismiss such arguments off the bat because we have noted that DONUD is independent from the problem of induction. Whether or not DONUD leads to other kinds of unreasonable skepticism remains to be seen.

How to gain knowledge?

My proposed anti-realism stance relies on a form of empiricism. I take as primary our experience in the sense that I take what our senses tell us for granted *to a large degree*. In accepting this I am fully aware that there are examples of sense experience which seem illusionary. There is the example of the straw that we put in a glass of clear liquid and we see that it bends, yet we seem to know it does not bend; our fingers do not hurt when we put them in water, neither is the straw bended when we pull it out. In response to this example and others of its kind, I claim firstly that the majority of our experiences is *not* ambiguous in such a way, or at least there is no reason to

[52] P. K. Stanford makes a similar remark in his "Refusing the Devil's Bargain", p.S3

doubt this overmuch. Just look around you in the room you are currently in; there is already a large amount of objects of which the reality is not so ambiguous. Secondly I claim that *a priori* there is *no such thing* as knowledge about whether or not the straw *really* bends. Perhaps the water bends all objects equally, similarly to how gravity bends space according to general relativity theory. We would then be unable to determine whether or not the straw really bends. For instance we would feel nothing remarkable when we put our finger in the water.[53] This leaves the reality of *straw-bending* underdetermined. But the –observable- facts are that we see a bend, and that we experience no awkward rearrangement of our physique when entering bodies of water.

That said, my stance should be slightly *surprising*, since it would entail a kind of theory-ladenness of observation that on first face renders moot the difference between underdetermination on the unobservable level and underdetermination on the observable level. How would we know our experience *can* be taken for granted? In response, I want to make a move towards a slightly Kantian philosophy. What science *should be* about, and what I think is the only important thing science *can be about* as I will try to argue, is the world as *we* experience it. This does not mean that theoretical knowledge is *right out*, it merely means that science should take our experience for granted and build upon that. If there is a method of building upon experience to reach theoretical truths *we can be realists.*

Keep in mind that when I talk about 'experience' I want to explicitly move away from just observation as if it exists independently from our mind and thoughts, and accept a degree of theory-ladenness in observation. This degree of theory-ladenness is not significantly different from the separation between *dinge an sich* and *dinge für uns* by Kant. Experience is not *just* our raw observation, but the way we interpret this observation in terms of the concepts we use to describe it. For instance we communicate using concepts of objects like *cars*, *chairs* and *cups of coffee*. We seem to automatically view the world in terms of these concepts. Philosophically speaking there is nothing in the observation itself which *forces* these concepts upon us. For example, we could think of the coffee cup as part of the coffee machine, and have one name for it: 'the coffee-cup-and-machine'. But in our experience the cup and the machine would be part of the same thing, and we would not refer to the coffee cup as something separate. A similar idea can be found in Goodman's Riddle. The idea is that the concept of 'green' could just as well have been 'grue' as long as grueness means that green turns to blue on some yet to be observed time. 'Green' does not force itself upon us by observation of past 'greenness' alone.

[53] This coincides with experience.

According to modern physics, objects like cars and cups of coffee are just collections of molecules. Where does the one object end and the other begin when the world is a collection of molecules? We cannot just say 'well those molecules belong to the chair and those to the ground it stands on' because this already assumes our experience in terms of concepts like chairs. A realistic interpretation of science would teach us that the molecules are the *real nature* of the world around us, and that how we interact with them results in our experience of the chair. But the chair is not real *an sich*, it is just a result of interactions of charged particles with electric fields of other charged particles. The chair doesn't really exist as we experience it; the reality of the chair is not a big continuous block of matter in a certain shape, the reality is that it is a bunch of molecules holding each other together – whatever that may mean according to science – with large amounts of space without matter in it.

I want to propose the following: there is no a priori reason to assume molecules are the real nature of the world around us, nor any other unobservable entity or concept. What is *real to us* is the world in terms of the concepts in which we view it. Science can speak about concepts and entities like molecules and forces in terms of conceptual frameworks to deepen our understanding of the way we experience the world, and even to create new experiences, but this does not necessarily imply that we have enough reason to believe such entities and concepts are real. What *is* real to us is our experience. The burden of proof thus lies on science –in particular the realist- to show that the world really is made up of molecules. Again, I'm positive towards any arguments that provide good reason to believe science. I am however skeptical about any position that takes it for granted that science can bring truth. Intuitively speaking science has an impressive track record so why should we consider science to have the burden of proof? I will return to this kind of argument later on, it is similar to the 'no-miracle' argument, and show that this argument itself relies on taking it for granted that science can bring truth.

The question that naturally comes up is how I am justified in taking those concepts for granted if they are about observables and how I deny that there is enough reason to believe in them as real if they are about unobservables. This is where underdetermination plays a role. Logically speaking the concepts that describe the world on the observable level are also underdetermined, that is they are theoretical to a degree[54], and we need a convincing argument outside of purely logical realms to pin the 'real' concepts. This argument is given if we make the move to consider our experience primary. That is, there is a reason to believe as true the concepts which we use to describe the observable world precisely because it is how we experience the world. We experience the world in

[54] See my previous discussion about 'the chair'

terms of *cars*, *cups of coffee* and so on, and since science should be about the world we experience, there is enough reason to believe these concepts are real in that sense. We have a priori no other reality to fall back on. For all intents and purposes these concepts are true, and they are true by virtue of being descriptive, and nothing else. The concept of 'green' just corresponds to something we have experienced and have given a name, but by virtue of this calling something green when it fits the experiences we have given the name green trivially makes it true.

We do not experience the world in terms of unobservable entities, however, because to experience them we need to have some way to reach them with our senses, so the argument cannot be repeated without further justification. In another way we can say that whatever the real nature is behind our experience, it does not matter *to our experience*. Whether there are electrons and protons and electric fields, or that matter is the continuous blob we experience and nothing else, should be irrelevant to our experience. We were experiencing matter as a continuous blob, or however we experienced it, before we knew about protons and electric fields, and gaining this knowledge did not change our experience. Underdetermination guarantees this, at least in the form of DONUD.

Like in the observable case, we need something to pin the unobservable concepts down. If I can show that there is no way to avoid underdetermination on the unobservable level, I show how there is nothing in our philosophical toolset that can pin these concepts down because other concepts are then always compatible with the same experience. This would mean that even if we accept that there is an unobservable world in which it is assumed that certain concepts are true we wouldn't know *which* of the possible concepts would be true, out of the set of applicable concepts known to us.

I am not unfair towards realism, and am willing to accept reason as a solution to the problem. After all, I rely on the reasonability of accepting experience as primary, and admit that there is no necessity for accepting this point of view. All I say is that it seems reasonable to accept our experience as true because it is real to us. There could be some argument, some form of reasoning perhaps, that allows us to pick one of the possible unobservable concepts and determine that we have enough reason to believe it as true even if we only take experience as primary, and that would be enough to allow us to defeat this nefarious underdetermination problem and to believe in unobservable entities as well, thus making realism acceptable.

As a final remark in this paragraph, let me make clear that my step to accept experience as primary is not a step which invalidates realism. I do not automatically reject realism by adopting the foundation of empiricism in this way. The arguments for DONUD and against realism stand with or without my empiricism. I have introduced this form of empiricism to combat the argument that underdetermination is equally a problem for the observable level as it is on the unobservable level. I claim that is not equally a problem, because the concepts we use on the observable level are pinned by our experience, while the concepts we use for the unobservable level are *a priori* not. If there is an argument which invalidates this distinction, which remains to be seen, anti-realism of this kind would collapse into skepticism of an unacceptable nature. Then we could no longer hold that statements like "the grass I see outside is green" are objectively true. Realism is still a valid alternative in that case, because it does not have this form of skepticism about either the observable or unobservable world.

My point of view above should be acceptable to the realist and anti-realist alike, the point being that realists will accept it, and add additional sources of knowledge or claim that there are reasons why this kind of empiricism should lead to realism. I have not said that DONUD and my empiricism prevent that there is a solution to DONUD. Such solutions will have to be spelt out and investigated, which will be the subject of the rest of my thesis.

What kind of anti-realism do I have in mind?

A. Kukla points out an interesting ambiguity in the interpretation of realism, which I will shortly discuss to make clear what sort of anti-realism I think is supported by DONUD. A standard interpretation of realism is "to believe in theoretical entities";[55] it can be interpreted as believing in the existence of theoretical entities even if we do not know anything about them, and it can be interpreted as believing that there are theoretical entities X and that these are real, including all their properties as professed by the theory that governs X. These kind of comments point to an interesting subtlety in the realism debate, namely that depending on your view on these matters, some realist positions could be interpreted as anti-realist positions. For instance, I feel sympathy for the first position, namely that there probably are unobservable entities. But I would consider such a position anti-realist if the point would be that our science does not provide justification for believing in the truth of any theoretical statements we have. Kukla names this first interpretation abstract realism.[56] In a sense my position is then a position of abstract realism; it seems likely to me that there are unobservable entities, I do not dare take the arrogant position that all we see is all there is. The point is that we cannot know whether or not science gives us true knowledge about these unobservables and whether or science is actually about the really existing unobservables. To reject abstract realism

[55] Kukla 1996, p. 140
[56] Kukla 1996, p.140

would imply that we do not believe in the existence of theoretical entities. This position seems to me to be going too far. The point is that we do not *know* anything about possible unobservable entities, in particular we do not know whether they exist or not. A truly sceptical position would not reject abstract realism. I would however still call myself an anti-realist.

Kukla mentions different forms of realism at the end of his paper which may escape his anti-realist arguments:

"But it still leaves both forms of feeble realism in the running – the rational warrantability of ascribing of nonzero probabilities to (a) some hypothesis of the form "theoretical entity X exists" (feeble concrete realism), or to (b) the hypothesis that there are theoretical entities (feeble abstract realism)."[57]

To make things clear, I would ascribe myself to a strong form of feeble abstract realism - I think there is a strong probability that some theoretical entities exist - and to feeble concrete realism – I think there is a possibility all our proposed current theoretical entities exist. However, I do not think these positions are properly called realism, even with the qualifier *feeble*; nonzero probability of a theoretical entity existing is *not* equivalent to having justification for believing in them. Perhaps high probability would be enough justification, but that is a discussion on its own. Accepting the possibility of theoretical entities existing, even those that our *current* theories suggest, seems to me to be par for the course for a sceptical attitude.

The observable/unobservable distinction

Before we head off into the marshes of the realism debate, we should address a question that arises from the discussion above. I seem to be taking for granted that there is a significant difference between observable and unobservable entities. In fact whether or not there is a valid distinction between the unobservable and observable world has been a point of discussion in the realism debate. In the following paragraph I wish to discuss the points raised against the distinction, and show how I am justifiably 'taking it for granted', based on the fact that I take experience[58] for granted, for which I have given – I think - good reason.

B. van Fraassen has written some influential papers and books in which he pays attention to the difficulties of the observable/unobservable distinction. His *constructive empiricism* may at first glance seem similar to what I am proposing, but there are significant differences. One thing that our views do have in common is a distinction between observables and unobservables. Particularly, van Fraassen says that 'science aims to give us theories which are empirically

[57] Kukla 1996, p. 162
[58] Note: this means not just pure observation, but the concepts we use to understand and communicate these observations as explained above

adequate; and acceptance of a theory involves as belief only that it is empirically adequate' and that 'to accept a theory is to believe that it is empirically adequate – that what the theory says about what is observable is true.'[59] Constructive empiricism implies being agnostic about what a theory says about the unobservable world, or at least *not believing* that what a theory says about the unobservable world is true.

A rejection of the distinction between observable and unobservable is an objection to constructive empiricism as well as to my own view. I will look at the literature involving van Fraassen, and see if we can learn a good way to enforce the distinction in my proposed view. In his 1980 classic, van Fraassen already responds to objections against the distinction by G. Maxwell. These objections were made against logical positivism, which had a similar distinction in terms of observation language and theoretical language. Maxwell rejects the distinction based on a rejection of all the arguments for the distinction. He argues that there is a *continuity* between observable and unobservable which means the distinction has no ontological status.

It is interesting that Maxwell feels that '[his] paper *should* turn out to be a demolition of straw men'[60] while his construal of anti-realist positions can now be deemed a straw man. We need to keep in mind, however, that Maxwell responded to admittedly flawed anti-realist positions of the time. He introduces a fictive story in which a scientist named Jones hypothesizes about unobservable organisms called 'crobes' which transmit diseases among humans. Jones' hypothesis is successful in the sense that as a result of taking measures against transmission of crobes, as proposed by Jones, diseases are transmitted less. He then introduces a few philosophical attitudes towards crobes, believing that they are real amongst others. A new instrument is invented which is said to allow one to observe crobes. The previously concocted philosophical attitudes towards the existence of crobes are now required to respond to this new information. Maxwell states of one of these attitudes: 'a more radical contention was that the crobes were not observed at all.'[61] It should be noted that my position is that we do not have enough reason to say we *know* we observed crobes. This is a much weaker position than Maxwell attributes to skeptics about the reality of unobservables. Most strikingly, he considers using instruments to be observation just as observing with the naked eye would be. DUD should have been enough to dispel such a claim, but he does raise an interesting point that I think is worthy of response.

How can it be, he says, that we can look through a pair of glasses and trust these as reliable sources of knowledge, but that we cannot look through a

[59] Van Fraassen 1980, p. 12 and 18
[60] Maxwell 1998, p.1052
[61] Maxwell 1997, p.1055

microscope and do the same? Isn't there a continuity of observation using instruments, the microscope merely being a strong version of a set of glasses? My answer is an emphatic *NO*. There are several reasons why we shouldn't believe there is a continuity, and at least one criterion can be formulated that explicitly breaks the continuity. It boils down to this: a pair of glasses is a different thing from a microscope!

Firstly, I echo Bas van Fraassen's critique that the observable (or not) nature of a physical entity should be considered from the point of view of man, or more precisely the scientific community since we are talking about science.[62] Things are observable (or not) *to us*, and not *in principle* or *ontologically*. *Having* to use an instrument is because something *is unobservable*. However, Maxwell's comparison of microscopes with glasses is still an issue for this point of view. Van Fraassen has a counter to this point by stating that all this continuity example does is show that observability is a *vague predicate*.[63] That means there is logically no way to formulate a precise point at which observability changes into unobservability, but does not mean there is no difference between observability and unobservability.

This is my second point; I think a careful look at how microscopes work will give us a broader understanding of the reason why anti-realists think that instruments are problematic epistemologically. Maxwell presumes that there is continuity between glasses and microscopes, I suspect because of his belief in a theory of light in which light is refracted through lenses. An obvious objection is that Maxwell is begging the question here. Anti-realists should *reject* a priori knowledge of how glasses and microscopes work in terms of refraction and thus that they work similarly. All we have, until we have reason to believe otherwise, is our experience that glasses make visible things 'appear sharper', for lack of a better word, and magnifying glasses make things bigger, and microscopes, of low enough power to make the comparison between an observable object when under and when not under the microscope, make things even bigger. We can say this safely as long as we can compare the resulting image with our naked eyesight.

This is why we can say that glasses reliably make us see things better. To say that what we see through glasses is the same as what we see without them, though less sharp, may at first glance seem to be a theory-laden statement. But it becomes reliable knowledge when we *experience* that the images are comparable. The same thing goes for magnifying glasses and low power microscopes. For the latter, imagine we have a magnification of x1 when looking at some hair. Now we increase the magnification steadily, and the hair appears bigger. We might even begin to see a scale-like appearance of the surface of the

---

[62] Van Fraassen 1980, p. 17
[63] Van Fraassen 1980, p. 16

hair. So far so good, but this is still not the domain typical of the entities the realism debate is about.

An interesting break in continuity happens at some point here. In an article called 'Do We See through a Microscope?' I. Hacking points out that high power microscopes do in fact *not* work the same way as do magnifying glasses. While the latter use laws of *refraction*, those microscopes use laws of *diffraction*. Theoretically, magnifying glasses and low power microscopes, and perhaps even the eye, are entirely different instruments from high power microscopes. Hacking concludes: 'I think that means that we do not see, in any ordinary sense of the word, with a microscope.'[64] So Maxwell's continuity does not hold at all.

But there is another way in which the continuity does not hold, and it does not hold for the same reasons DONUD should be taken seriously[65]. Image once again we zoom in on the hair, assuming that our microscope has no trouble doing this by utilizing the laws of refraction.[66] We see the scale-like appearance of a hair appear when we zoom in on it. We have enough reason to believe that the hair really does have some scale-like properties, whatever that means in observable terms, because we know it is the hair, since we zoomed in on it. The hair is an observable object, and we can compare the largely magnified image with our plain eyesight observation of the hair. I see no problem there.

Now imagine we see something which we have not seen before, indeed as it turns out what we are seeing is too small to see with the naked eye.[67] Assume realists would call these 'microbes' or 'bacteria' or some other microscopic creature. Does this observation through the instrument validate our belief in the existence of microbes, or whatever the creature at hand would be? At first glance, because of the analogy with the hair, it might seem that yes, we do have enough reason to believe. But the analogy does not hold. We have *nothing* to compare the image with in our reality. Words like 'microbe' or 'bacteria' are *theory-laden* terms; they play a role in our wealth of theories and hypotheses so these terms have 'baggage' so to speak. This means that if we give these names to the observed *something*, we assume that it has all kinds of properties not obvious to us even with the use of the same microscope.

It is interesting that Maxwell addresses the point of 'unobservable in principle' by stating that our capacity to observe might change at any moment because of various amounts of reasons. He uses an example in which a drug is invented which 'vastly alters the human perceptual apparatus-perhaps even activates latent capacities so that a new sense modality emerges.'[68] Bas van

---

[64] Hacking 1985, p. 133
[65] On the level of unobservables…
[66] Or, for strict empiricists, that the microscope has a zoom level at which obviously it shows things comparative to what we see without the instrument.
[67] In other words, the object appears when zooming in and disappears when zooming out.
[68] Maxwell 1998, p. 1058

Fraassen has pointed out the obvious flaw in this reasoning[69] and I wish only to add that, indeed, if such drugs were available and *reliable*, which is a different matter in itself, I would applaud such a find as a major improvement in science.

However, the argument can be turned and used against Maxwell's point. His claim comes down to accepting that use of instruments is as reliable as using the naked eye since human observation is in principle no different from using instruments. Maxwell concludes 'that our drawing of the observational-theoretical line at any given point [...] has no ontological significance whatever.'[70] If we would take sensory enhancement drugs the result would be no different from peering through a microscope. Both merely enhance our range of observation. I disagree for the reasons already discussed above, but I think there is an interesting analogy between hallucination-inducing drugs like LSD and instruments to be salvaged. Isn't the point of such drugs that we are convinced hallucinations aren't real in the physical world *precisely because* we do not experience the content of them normally –in other words, with our naked eye-? If we are skeptical about the reliability of *these* drugs, why are we not a priori skeptical about instruments for the same reason?

Having rejected that there is in principle no observable/unobservable distinction, there still is a problem having to do with determining whether or not something is in fact observable. Alan Musgrave wrote an objection to van Fraassen's use of the distinction, which I think should be discussed before accepting it. Musgrave's objection is that, given van Fraassen's 'rough guide' to determine what is observable, there is an incoherence in accepting what a theory tells us on the unobservable level and believing as true what a theory tells us on the observable level. Van Fraassen's 'rough guide' is as follows:

"X is observable if there are circumstances which are such that, if X is present to us under those circumstances, then we observe it."[71]

The crux is that van Fraassen's central attitude towards statements of an accepted theory allows *theory* to tell us what is observable, namely that we can believe as true what a theory tells us about the observable world, in particular that an object is observable; that an object is present to us under certain circumstances we observe it. Note that van Fraassen does not say observability is theory-dependent, but that it is a matter of practical consideration. For example, van Fraassen would recognize that since we cannot observe an electron, we must rely on theory to tell us its properties –if it exists-, including the properties concerning size and visibility. The problem is that in constructive empiricism you can only believe in what a theory tells us on the observable level.

[69] Van Fraassen 1980, p. 16-17
[70] Maxwell 1998, p. 1062
[71] Van Fraassen 1980, p.16

A theory can tell us that X is observable according to the rough guide, and we have enough reason to believe this is true.

Now we want to be accepting -as empirically adequate- but not believing -as true- about unobservables. As Musgrave points out, constructive empiricists 'cannot believe it to be true that *anything* is unobservable by humans'[72] since any statement 'X is unobservable' is not a statement about an observable. But if we cannot know it to be true that certain objects are unobservable, then the statement that we *must* be agnostic about any statement about unobservables is superfluous; we cannot know whether or not such statements refer to anything in the real world in the first place.

This is not only a serious objection to constructive empiricism, but it seems to be a serious objection to any form of anti-realism which rejects knowledge about unobservables but not about observables. This includes my own form of anti-realism, so I think it is time well spent to discuss Musgrave's objection. Van Fraassen has written a curious response to Musgrave. He makes the claim that if T entails the statement "X is unobservable" but it is actually true that X is real and observable, then T is not empirically adequate, since otherwise it would state something about the observable which is not true; namely that X, which is observable, is unobservable. So, according to van Fraassen, as long as a constructive empiricist believes T which entails the statement "X is unobservable" to be empirically adequate, he believes it to be true that "X is unobservable" if X is real, on pain of contradiction.[73]

But this seems to only reinforce Musgrave's point. Van Fraassen has now admitted to believe as true something about the unobservable, if it is real, namely that "X is unobservable", while this is forbidden in constructive empiricism. Van Fraassen even gives a *procedure* of determining something about unobservables, namely that they are unobservable. I think the problem here lies in van Fraassen's overly positive acceptance of theory as the harbinger of what is true about the observable world, which is the basic epistemic attitude of constructive empiricism. I will elaborate on this after I have discussed F.A. Muller's explication of van Fraassen's rough guide and Musgrave's objection. Muller tackles the issue in a way that illuminates the problem in the attitude of constructive empiricists and shows the importance of Musgrave's objection.

It is worth pointing out as Muller does that Musgrave's objection against the observable/unobservable distinction is not criticism about the distinction itself, 'but against *drawing it within the confines of CE*'[74], and my version of anti-realism suffers the same fate. Muller introduces Musgrave's argument by pouring it into the mould of modal logic. I will try to paraphrase his results to

---

[72] Musgrave 1985, p. 208
[73] Van Fraassen 1985, p.256
[74] Muller 2004, p. 638

avoid overly technical accounts of what can be said in more straight-forward terms. First we make clear that any statement about a physical object is empirical *iff* the object is real and unambiguously observable. So if a statement about a physical object is *not* empirical, the object could either be unreal *or* unobservable *or* both. To illustrate, an example of an unreal and observable object is a unicorn, and an example of an unreal and unobservable object is the Snigg's boson, the Higg's boson's brother from another mother. So if a theory is empirically adequate and it entails a statement which is empirical, then according to constructive empiricism we have reason to believe in the truth of that statement. Conversely, if a theory is empirically adequate and it entails a statement which is not empirical, then according to constructive empiricism we have reason to only accept the statement, but not to believe that it is true. Not believing in the truth of a statement is not the same as believing the statement is false, constructive empiricism typically promotes agnosticism about accepted statements for which we do not have enough reason to believe they are true.

If a statement is about unobservable aspects of the real world, then it follows that the statement is not empirical, and that we therefore have no reason to believe it as true. It should be noted that it does not imply either that we have reason to believe the statement is false. A statement of the form "X is unobservable" is a statement about an unobservable aspect of the real world. We have no reason to believe it is true. Thus, it is impossible for constructive empiricism to draw the distinction between observable and unobservable in a meaningful way, because we cannot be certain about the unobservable part of the distinction. It might be objected that being agnostic about a physical object being unobservable is enough, because we are agnostic about any statement about the unobservable. But this is to say that we must be agnostic *about part of the core attitude of constructive empiricism*, namely that we must be agnostic about the unobservable when we can never know whether or not there are unobservables; what is there to be agnostic about? We still have the intuitive concept of unobservable which might seem viable, but from the point of view of constructive empiricism it is impossible to implement it in a coherent way, because we will never be able to determine that a statement really is about unobservable aspects of the real world.

As Muller further points out, Musgrave's criticism can even be extended to unreal but (obviously) observable objects, like Pegasus, Hydra and Cyclops. Because they are unreal, or at least so we believe, any statement about them is not empirical. But we should be agnostic about non-empirical statements. So we cannot conclude whether or not objects of fiction which are intended to be observable are in fact observable.

Let's go back to van Fraassen's reply to Musgrave. According to Muller's analysis this reply can be reconstructed as follows: if any statement about an electron is a non-empirical statement, and these statements belong to an empirically adequate theory, then electrons must either be unobservable or unreal. If electrons are real, then they must be unobservable. This means that constructive empiricism 'can believe that electrons are unobservable-if-they-exist'[75]. But this is exactly the problem, because part of the core epistemological attitude of constructive empiricism is to only believe in the truth of a statement if it is about observable aspects of the real world. Believing that electrons are 'unobservable-if-they-exist' is a violation of this attitude.

It might be objected, as Muller notes, that constructive empiricism does not rely on theory to tell us what is observable. After all van Fraassen points to the anthropocentric nature of 'observable':

'The human organism is, from the point of view of physics, a certain kind of measuring apparatus. As such it has certain inherent limitations – which will be described in detail in the final physics and biology. It is these limitations to which the 'able' in 'observable' refers – our limitations, qua human beings.'[76]

Without relying on theory to tell us what is observable, Musgrave's objection seems defeated, since it relies on theoretical statements of the form "theory T has statement 'X is unobservable'". But it turns out that the argument can be repeated. Even if we let experience dictate what is observable, it *cannot dictate what is unobservable* within the confines of the epistemological attitude of constructive empiricism. If we come to the conclusion that we aren't capable of observing an object, we can believe that either the object is unobservable and exists or that the object does not exist. But we cannot believe either one of these options without knowing about the other; we must know an object is real before we brand it unobservable, and vice versa. A solution would be to accept a third option: If we come to the conclusion that we aren't capable of observing an object, we can believe that the object is unobservable-or-unreal-or-both. But this prevents us from believing that Pegasus is observable and that we are therefore justified in believing the nonexistence of Pegasus, but van Fraassen *wants* to believe that Pegasus is not real. If we do accept the third option, a revision of the observable/unobservable distinction follows naturally, as we will see below.

Muller claims that nothing less than an extension of the epistemic policy of constructive empiricism is required. Without an extended policy constructive empiricism cannot believe anything about the unobservable, not even that it is unobservable or that it exists. Van Fraassen ultimately agrees that this criticism is valid, for in a paper written together with Muller, he adds the following rule to the epistemic policy:

[75] Muller 2004, p. 648
[76] Van Fraassen 1980, p. 17

"0. If you accept T, and Y is (un)observable according to T, then believe so."[77]

According to van Fraassen and Muller, this is 'perfectly compatible with the spirit of constructive empiricism'. However, this amended policy seems merely a trick to me, for it fully violates the original stance of constructive empiricism in the same way Musgrave pointed out. We are supposed to be agnostic about anything a theory says about the unobservable, but *magically* we can be sure about the theory when it tells us a physical object is unobservable. Van Fraassen and Muller give *no reason* to make this amended stance credible, indeed they elevate it to an axiom, and it even marks a great retreat from the original view of observable being an anthropocentric concept.

Even if this amended policy would be an answer for constructive empiricism, it is not for my version of anti-realism. According to my anti-realism, this new epistemic policy runs into the same problems of underdetermination as other statements about unobservables, and Musgrave's problem must be solved by turning to another epistemic policy. Muller has proposed a different epistemic policy based on experimental findings; a solution more in tune with my version of anti-realism. Muller proposes that we look at experimental findings in order to tell us the limits of human perception. This in turn will be a guide to determine what is observable and what is unobservable. The results of his inquiry into experimental results of the physics of man are:

'The sensitivity-threshold $s : \lambda \rightarrow s(\lambda)$ *informs us how much energy per second is needed for the retina to send a signal to the brain and consequently for us to see some thing.'*

'***Scientific Criterion*** *[for being observable]. On the presupposition there is temperature* T *and a pressure where object* X *and* $p \in \mathcal{E}$ *survive,* $\mathrm{Obs}(X, \mathcal{E})$ *iff*

$\exists d \in [10\text{cm}, R_{@}], \exists \lambda \in [400, 800]\ \text{nm} : s(\lambda) < E(T, \lambda, S_X, d)$

Where $R_@$ *is the radius of the universe (about 156 billion light-years),* $s: \lambda \rightarrow s(\lambda)$ *is the sensitivity-threshold of the human eye, and* $E(T, \lambda, S_X, d)$ *is the total energy of emitted-cum-reflected light of wavelength* λ *by object* X *having surface* S_X *at distance* d *from* $p \in \mathcal{E}$. *If there is no such common survival temperature and pressure for* X *and* $\mathcal{E}$, *we consider* $\mathrm{Obs}(X, \mathcal{E})$ *to be indeterminate; and if there is but* X *transmits all visible light and neither reflects nor emits any (such as the invisible man) , then we call* X *unobservable.'*[78]

And an updated verson of the 'rough guide':

'$\mathrm{Obs}(X, \mathcal{E}, \mathbf{L})$ *iff* $\forall p \in \mathcal{E}, \exists \mathcal{M} \in \mathbf{L}: \mathrm{tr}(\mathcal{M}, \mathrm{Front}(p, X) \wedge \mathrm{Sees}(p, X))$'[79]

Where **L** is the wave theory of light, $\mathcal{M}$ is a model of **L**, $\mathrm{tr}(\mathcal{M}, B)$ means that B is true for that model, Front(p, X) means that X is at rest and in front of p which has healthy eyes open and Sees(p, X) means that p sees X veridically. Roughly

[77] Muller & van Fraassen 2008, p. 204

[78] Muller 2005, p. 81-82. $\mathcal{E}$ is the epistemic community and $p \in \mathcal{E}$ is a member of the epistemic community.

[79] Muller 2005, p.83

this says that we can always imagine some world in which the wave theory of light is true and object X is seen by p. Muller gives some examples to reinforce this definition, and I will repeat some here as they are educating.

A living Tyrannosaurus is observable, since we can imagine a world compatible with the wave theory of light in which a Tyrannosaurus lives and walks in front of us and we see it. As Muller points out such an imaginary world *has* been created in the movie 'Jurassic Park'. Electrons are unobservable since for every imaginary world in which the wave theory of light is true, electro-magnetic waves ignore electrons like "a tidal wave 'does not see' a grain of sand."[80] Black holes are unobservable because according to the wave theory of light electro-magnetic waves cannot reach us from a black hole. A fish living only in the deep black sea where we ordinarily cannot see is observable, because we can imagine a world in which they swim in an aquarium in front of a member of the epistemic community.

At first glance an obvious counter is available: this new guide does not avoid Musgrave's problem in the same way the solution given by van Fraassen and Muller above. It does go one step beyond that solution, however. It does not merely take for granted that a theory says something is unobservable, because theories themselves *do not say* whether or not something is observable or not. They only give certain properties of a physical object, and it depends on other theories, like the theory of light as Muller proposed, whether or not these properties render an object unobservable. However, it is still dependent on theory, and grants the wave theory of light some sort of privileged status. Muller claims that while van Fraassen might be in trouble, the same trouble does not befall his own criterion.[81] Most importantly, he claims that 'CE [that is, the adjusted version of constructive empiricism] assigns a privileged status of sorts to theory **L**, but that is a direct consequence of the privileged epistemic status CE assigns to actual observables, in good empiricist tradition.'[82]

I do not see how even CE can assign a privileged status to theory L as far as it says things about unobservables, without violating CE's central epistemic attitude about statements about the unobservable. To do so would be to fall into the same trap van Fraassen's and Muller's solution above does. It assigns some sort of 'magic' status to the wave theory of light as far as determining whether or not something is unobservable goes, without further justification. In particular, it seems **L** is as guilty of the 'inflationary metaphysics' constructive empiricism wants to avoid as any other theory. From the sceptical view point of the anti-realist this raises the question: How can we be so sure the wave theory of light, a

[80] Muller 2005, p. 83
[81] Muller 2005, p. 85
[82] Muller 2005, p. 85

theory ostensibly about unobservable things called 'waves of light', provides us with a true description of the real world on the unobservable level?

There is something to be salvaged from Muller's suggestion. Why not focus our attention to the purely observational results of the experiments he has been looking at? But different from Muller's approach, we would do so without invoking the wave theory of light. The result would be a simple list of objects we were able to observe and their measurable properties, as far as they could be measured with the naked eye. For instance: 'Under green light an object of about 500 micrometers could be observed by the sharpest observer, but somewhere around that size lies the boundary.' In the same way we could expand our concept of observable to include all senses, in order to encompass all our experience.

True, we are still not capable of determining the properties of the unobservable. But that is no problem for my version of anti-realism, and it should not be a problem for constructive empiricism as long as it is adjusted in its core epistemic attitude. The problem as I see it is that van Fraassen proposes on the one hand an anthropocentric 'rough guide' of what the property 'observable' means, but on the other hands wields an observable/unobservable distinction which is ontological. To make matters worse, this distinction says nothing about epistemological matters; it does not connect in any way with how we learn that something is really observable or not. Van Fraassen's distinction only functions in a case where we have complete empirically adequate theories, in the sense that we have categorized all observable events. In actual science, which is ostensibly open-minded, the empirical adequacy of theories is always controversial.

I want to suggest the following revision of the epistemic attitude based on the discussion above: we should believe as true anything an accepted theory says about what we have determined to be observable, and keep an agnostic attitude towards anything a theory says about the not-yet-determined-to-be-observable, let's call this the *nydobservable*[83]. This means we fully embrace the third epistemological attitude towards 'unobservables' discussed above; electrons are 'unobservable-or-unreal-or-both' or nydobservable. This attitude accepts that perhaps there are observable objects which have escaped our notice –our theories are not complete-, but we should place our trust in science, that if we keep continuing our quest for the expansion of empirical knowledge such objects which have escaped our notice will come to light. So this attitude is even applicable to theories *as they are*, not merely theories in the *never-never land* of complete empirical adequacy. Furthermore it embraces revision and criticism, both based squarely on experience.

[83] Pronounced *neet*-observable.

I think the discussion throughout this section warrants such a revision of the epistemic attitude of constructive empiricism, and I think it is a good point to have learned from all the discussion about the observable/unobservable distinction. Thus, there is *no meaningful observable/unobservable distinction for purposes of anti-realism about unobservables*, precisely because our attitude towards things we cannot observe should be agnostic -or even believing-as-false- according to such a stance. A priori we know not whether or not a hypothesized object exists, much less do we know anything about any physical properties it might have. These physical properties include the point of discussion: whether or not it has such physical properties that it is unobservable to us. An object might be unobservable, unreal or it may have escaped our notice thus far if we abandon the criterion of complete empirical adequacy, but we should be agnostic about *anything* a theory says about these objects, simply because we have no method of verification due to underdetermination. This stance is fully compatible with the sceptical attitude of an anti-realist.[84]

On the other hand, objects can *gain* the observable status. This is a privileged status in which the object is acknowledged to be real and observable. A statement about such an observable object can be verified as long as it is a statement about observable properties, and an accepted empirically adequate theory can be trusted in what it says about these observable objects.

There *is a tenable distinction* between that which has gained the observable status and that which has not. This distinction avoids any of the problems discussed above, and in particular it solves Musgrave's problem. Following Musgrave, we are agnostic about what theories say about things which have not yet gained the observable status, including statements that an object is unobservable. In contrast with Musgrave and Muller, there is now no contradiction. We can believe that an object has not yet gained the observable status, while being agnostic about it being unobservable. It might be observable and escaped our notice, or it might be unreal, or real but unobservable, but we are not sure and maintain a safe sceptical attitude which avoids 'inflationary metaphysics'. This stance is not strictly compatible with constructive empiricism, but a revision could be made such that constructive empiricism does not maintain an observable/unobservable distinction but an observable/nydobservable distinction. I think such a revision is fully in the spirit of the original intent of constructive empiricism.

In conclusion, we have learned a lot from criticism about the observable/unobservable distinction, and in part that a lot of the criticism has a grain of truth, but that it ultimately does not dissuade from adopting anti-realism. Van Fraassen's constructive empiricism as presented is to be rejected

[84] Unless we of course find a way around underdetermination, a quest undertaken further on in this chapter.

on the same grounds as realism is to be rejected: it places an unjustified amount of faith in theory. Van Fraassen's attitude implies that one can rely on an accepted theory to predict something on the observable level even if we have never experienced anything like it. I reject such an attitude since novel predictions should also function as tests, and it strikes me as no surprise that it is ultimately incoherent as Musgrave pointed out. I am reminded of someone who put it to me that he puts his faith in Newton's laws every time he gets into his car; that it is going to do what he expects based on belief in those laws. I'd rather think it a healthy attitude to put faith in your experience instead of any Law; that your experience is that the car always works in the same way and that you put your faith in that it is going to work this time for the same reason. That, to me, seems a much more tangible reason for trusting one's car to work.

Introducing ampliative inference

We now move to arguments against underdetermination. This should be the main attraction of this thesis. Once again I will review some of the arguments proposed to render underdetermination unproblematic, in the hope of sharpening my point of view. These arguments range from including more than just the data itself, like values such as 'fruitfulness' or 'simplicity' which may or may not relate to the available data, or relying on other types of inference than deduction, like induction. What all have in common is a departure from deduction as the only valid method of inference. I want to explicate how exactly common criticism to arguments of underdetermination leads to the acceptance of more than deduction alone, and how we should gauge these arguments in the light of DONUD.

Larry Laudan is well known for his criticism of many points of view connected with relativism, particularly as proposed by the 'sociology of science', a research tradition which investigates the ways in which knowledge becomes socially accepted among the scientific community. He mentions that he has written a series of articles against epistemic relativism.[85] In an article called "Demystifying Underdetermination" he wants to deconstruct the underdetermination thesis so that it either is trivial *but harmless* or forceful *but not valid*. He claims that in the literature what is usually called 'the underdetermination thesis' actually consists of many different underdetermination theses and that mixing up these different theses under one name has caused non-credible forms of underdetermination to take hold with the community.

He introduces two forms of the underdetermination thesis:

[85] Laudan 1998a, p. 322

Humean underdetermination (HUD) – For any finite body of evidence, there are indefinitely many mutually contrary theories, each of which logically entails that evidence.[86]

Quinean underdetermination (QUD*) – Any theory can be reconciled with any recalcitrant evidence by making suitable adjustments in our other assumptions about nature.[87]

Laudan points out that these two are not equivalent in the least. The first only speaks about deductive logic, and the second speaks about 'reconciliation'. In any non-trivial version of QUD*,[88] reconciliation involves the question whether or not it is rational to accept a theory. But it is not a priori clear that deductive logic and reason for acceptance are equivalent. Laudan acknowledges that deduction alone is not enough to move from scientific data to theory acceptance, but states that there is more to theory acceptance than just deduction. This is something most philosophers of science seem to agree on; nearly everyone seems to have realized that pure deduction is severely limited when it comes to establishing scientific facts at the theoretical level, and in reality other reasons for holding a theory than purely deductive arguments are often involved.

For instance, we could imagine instead of a force of gravity, an invisible leprechaun with an invisible rope swings the planets round the sun. If we assume that there is no way for us to interact with the leprechaun and that it reproduces the movement of the planets relative to the sun as we know it, purely deductively this new 'theory' would be a valid alternative to Newton's force of gravity. We do not even fathom such a theory for different reasons. We can think of actual examples in the history of science which are more convincing than the invisible leprechaun situation. Einstein's special relativity and Lorentz' electron theory are a well-known example of empirically equivalent theories; the choice between believing either of them over the other could not have been made based on purely deductive grounds. The transition from the Ptolemaic model of the movement of heavenly bodies to the Copernican heliocentric model is famously based on arguments not of 'harmony' and 'simplicity' but not because the Copernican model is an empirically better model of the observed motions of heavenly bodies. Thus, it seems reasonable to accept other forms of inference which are just as valid as deduction.

Laudan states that there are ampliative forms of inference; these broaden the scope of deductive methods and add information in non-deductive ways. Note that I do not equate non-deduction with induction here. As said before, I want to be clear about the kind of inference I am talking about, and want to avoid the pitfall of equating induction –which kind?- with everything that is not deduction.

[86] Laudan 1998a, p. 323

[87] Laudan 1998a, p. 328, asterisk mine.

[88] Note: QUD* is actually a formulation of the problem of auxiliaries, in an extremely strong form. Not to be confused with QUD.

As mentioned before, let's use the term adduction for every kind of inference which aims to amplify deduction for scientific purposes. Adduction includes induction, but also includes appeals to concepts like simplicity, inner consistency, fruitfulness and so on.

I accept Laudan's criticism partly; I accept that deduction is not necessarily the only valid method of inference, but I want to explicate which ampliative methods are acceptable and which aren't and what they are acceptable for in the context of the realism debate. Doing this will allow me to argue in favour of anti-realism without running into what is called 'inductive skepticism', and without making the mistake Laudan points out; confusing different kinds of underdetermination.

The use of adduction is Laudan's argument against QUD*. Because, as Laudan states, ampliative rules of inference allow us to go beyond deduction; it is not true that any theory can be reconciled with any recalcitrant evidence *when considering adduction*. Universal reconciliation would only be possible if we allow deduction as the only valid method of inference. We cannot *reasonably* expect to reconcile any theory with any recalcitrant evidence. I think the invisible leprechaun example above is a good example of why this argument is credible. While the invisible leprechaun model can be made to agree with all observable facts, it is not reasonable to propose this as a genuine scientific theory, and it is not reasonable to accept it as a possible contender for progress.

If we follow my analysis of Quine's article in the previous chapter, we can raise the issue whether or not Laudan has an adequate response to Quine's linguistic argument.[89] Quine can reconcile any theory with any evidence because the meaning of all involved terms is not set. We could even change what it means to 'reconcile', or abandon logic or even 'plead hallucination' as Laudan refers to.[90] Laudan raises the point that even here Quine does not talk about whether such a step is reasonable or not, so Laudan's criticism is easily extended: can you *reasonably* change the meaning of terms? The question can still be raised whether or not what counts as reasonable should stay fixed. But as I am advocating DONUD *and not QUD* I will not try to defend QUD any further. DONUD should be enough and is a *logically weaker* thesis.

QUD is as stated a rather far going thesis, but DUD is much more restricted. How does it fare under Laudan's criticism?[91] Laudan mentions Duhem briefly; he considers the point of holism to be the most interesting of the points Quine raises.[92] According to Laudan, there is a version of DUD which he calls the 'compatibilist version' which allows only *dropping* auxiliary hypotheses in the face of disconfirming evidence in an experiment. This agrees with my reading of

[89] In other words, that QUD is not the same as QUD*.
[90] Laudan 1998a, p. 327
[91] And by extension, how does DONUD fare?
[92] Laudan 1998a, p. 328

DUD in so far as Duhem leaves it open what can happen to core or auxiliary hypotheses once they are confronted with disconfirming evidence. The whole set of hypotheses is confronted with the evidence as a whole, and falsified only as a whole. Where fault lies, and whether or not it lies with auxiliary hypotheses or broken machinery or an occurrence of an anomaly is open according to DUD, and such questions need to be answered by *le bon sans*. Note that DONUD allows *but does not necessitate* dropping auxiliaries.

There is also another version of holist underdetermination which Laudan calls the 'entailment version' which also allows *replacement* of auxiliaries with other auxiliaries. Laudan explicitly rejects the entailment version on the grounds that no support for the universal availability of such replacing auxiliaries is given, and that even if we grant this, there is no reason to believe the acceptance of such auxiliaries is reasonable or rational.

Grünbaum attributed to Duhem a point of view similar to the entailment version, which according to Laudan is not a good representation of Duhem's holism.[93] Duhem himself claims that falsification of an isolated hypothesis is impossible, but not that any falsification may be rectified by replacing auxiliaries. He certainly did say that *sometimes* such replacements are justified and possible, as witnessed in history. So far DUD seems safe from Laudan's criticism. Indeed, I think Laudan sees in Duhem's doctrine of the 'good sense' exactly his own point; that in the face of a failure of deductive logic to progress in science, scientists rely on good sense. And what other name for good sense is there but reasonability?

However, this is not enough for DONUD. DONUD requires at least a weak version of the entailment version. Let's recap the situation. Laudan's criticism is that it is not true that any theory can be reconciled with any recalcitrant evidence. I agree in so far as I agree with Duhem; even Duhem says that groups of hypotheses *as a whole* can be falsified. So if we regard a *group* of hypotheses[94] as a 'theory', QUD* is not supported and in need of revision. Even though I accept DONUD I do not want to support such a far going thesis. The only thing we really need for an underdetermination argument to be forceful is to make reasonably acceptable the claim that we cannot *pinpoint* the hypotheses which give a true description of the real world. In other words, that there is always at least a, possibly small, number of reasonable rivals available, with the provision that the community might currently not be aware of these potential rivals -*it is the principle of the thing*. But this claim is much weaker than QUD*. I think DONUD as given in part I provides ample reason to believe this is the case at this point, because given the presence of DONUD there will always be room to maneuver in subtle ways. The Lorentz-Einstein case is a good example: It is not

[93] See Laudan 1965
[94] A group that of course includes auxiliaries.

possible to determine the non-existence of the ether using the Michelson-Morley experiment, nor does it seem possible to conclusively argue that Lorentz' theory had unreasonable assumptions in response to the result of that experiment; Lorentz' explanation of the Lorentz-contraction was troublesome but it did have merit in the sense that it explains the contraction in terms of physical interaction, something which Einstein's theory lacks.

In addition, I think the burden of proof for showing that underdetermination is not a valid argument for anti-realism, at least considering DONUD, lies on the realist. They need to show how reasonability exactly can solve the underdetermination problem, and we should not a priori accept reason as a catch-all solution to underdetermination. From the point of view of the epistemologist, whose job it is to determine how we can justify scientific knowledge, we started knowing nothing, then we added experience as a baseline, trying to expand on it but quickly stumbling upon underdetermination. If scientific knowledge is to be justified, solutions to underdetermination need to be justified. While Laudan's argument on the face of it seems reasonable, I want to warn against accepting it all too readily based on some intuitive affinity with a realist interpretation of science. As it stands this intuitive realist interpretation is not justified; underdetermination is such an obstacle that a solution to it should be given. What *would* provide justification is giving a detailed account of a form of adductive inference, and so I will be looking at those for the remainder of this thesis.

As an aside I would suggest that the history of science provides an excellent track record for the entailment version as proposed by Laudan. In the practice of everyday science, think of what a scientists *actually does* in the laboratory, the failures of science vastly outweigh the successes of science. The publications of successful experiments scientists put in their journals are not a good indicator of this; articles are usually the final product of a process of trial and error in experimentation; only the successes.[95] A. Fine refers to 'the successful tips of the mountains of failures.'[96] We should not allow ourselves to be overwhelmed by the eventual successes of science – which are admittedly impressive – to forget that for each such success there are months or years of hard work and failure. In light of the fact that these scientists manage to save currently accepted theories from consistent failure in the laboratory, an important aspect of science seems to be the constant search for auxiliaries –sometimes even core hypotheses- to replace the suspected faulty ones. Certainly, this does not *logically* imply that for any auxiliary or hypothesis there is a replacement. Seeing the absolutely gigantic mountain of 'replacements' in the history of science, in

[95] Any experimental scientist will tell you that the actual process involved more error than success. Collins also gives at least two credible examples of actual laboratory work which also involve more error than success.
[96] Fine 1986, p.153

some respects science seems like nothing *but* endless tweaking on existing sets of hypotheses. It seems to me at least plausible in this light to assume that for almost every auxiliary or hypothesis there is a replacement. Think of Collins' TEA laser experiment; the copy of the Jumbo laser did not work despite being a copy made by someone who worked on the original. Many 'auxiliary hypotheses' were considered; wrong positions of capacitors, wrong polarity of a spark gap, changing distances between electrodes, damaged components, thicker or thinner wires and so on. The bottom line is this; scientists are in practice fairly capable of finding auxiliaries to replace suspected faulty ones. What, as far as I am concerned, is the more important matter is Laudan's original point: Are such replacements rationally acceptable?

This is the main point in Laudan's article: by appealing to extra-logical rationality there is a way out of the 'hold on come what may' scenario.[97] To the point made in the previous paragraph realists might interject that all the cases of replacement of auxiliary hypotheses, or even core hypotheses if necessary, in the history of science have in fact been reasonable replacements, in the sense that it is not true that *just any* replacement might do, but only those which have some sort of cumulative value. I would reply that this kind of hindsight judgment reeks of *Whig history*, the kind of history of science that *presumes* science is a progressive enterprise. This type of history has been rejected along with positivism for good reason; first and foremost because looking back we find undeniable cases of failure. Each of Collins' case studies shows the pivotal role of failure in actual research. It would assume an uncanny ability of scientists to pick out the right replacement each and every time a replacement hypothesis is put forward. Indeed, Laudan himself refers to the role 'false' hypotheses can play elsewhere[98]; science seems to be able to pick out replacements which are later deemed false and yet use of these false hypotheses can be highly successful in the mean time, convincing scientists of their truth.

What can we conclude from the discussion above? That criticism of the kind Laudan has about QUD* also applies to DONUD. DONUD is a problem for realism because we have not allowed much beyond pure deduction in our quest for justification, except for experience. Ampliative forms of inference, which I will call adduction, have been suggested as a solution. My response to this is that we cannot just accept certain forms of adduction, that they need to be spelt out and reviewed before we can deem them reasonable solutions. The burden of proof lies upon the realist to find justification for such solutions, but I will take up the gauntlet and take it upon myself to investigate potential adductive methods.

In addition, it turns out that DONUD does *not* require that any theory can be held *come what may.* While on the face of it DONUD suggest radical

[97] Laudan 1998a, p. 325
[98] See Laudan 1998b

underdetermination on the level of nydobservables, it does not necessarily imply total relativism. All which is required from my side is that I explain how DONUD may be able to reject outrageous forms of epistemic relativism, while maintaining the core thesis that we cannot pinpoint a single hypothesis from among a group of reasonably acceptable hypotheses as the true description of events we are investigating. Note that I have already given a defense against epistemic relativism on the level of observables.

I will now continue to investigate specific forms of adduction with the aim of showing how they fail –or fail to give good promise- to provide a satisfactory solution to the problem.

Accounts of support and confirmation

One of possible extra-logical factors playing a role in scientific decision making that Laudan points out is that of empirical *support*. Empirical support stays close to "the facts" because it employs nothing other than them. It does ostensibly not require 'extra factors' like simplicity or consistency. It does require more than pure deduction based on the facts, namely some kind of adductive approach to the facts, which then leads to the notion of support as distinct from deducible empirical consequence. But support requires *facts*. Realists and inductivists seem content to use the word *fact* for all kinds of unobservable processes or objects which are subject to the criticism I have been putting forward. I will use the term fact only for that which has been shown true beyond doubt. A fairly well known example is that of the particle in a cloud chamber penned up by van Fraassen:

"The theory says that if a charged particle traverses a [cloud] chamber filled with saturated vapour, some atoms in the neighbourhood of its path are ionized. If this vapour is decompressed, and hence becomes supersaturated, it condenses in droplets on the ions, thus marking the path of the particle. The resulting silver-grey line is similar … to the vapour trail left in the sky when a jet passes. Suppose I point to such a trail and say: 'Look, there is a jet!'; might you not say 'I see the vapour trail, but where is the jet?' Then I would answer: 'Look just a bit ahead of the trail … there! Do you see it?' Now, in the case of the cloud chamber this response is not possible. So while the particle is detected by means of the cloud chamber, and the detection is based on observation, it is clearly not a case of the particle's being observed."[99]

In the realism debate we sometimes find, in contrast to the position van Fraassen takes here, reference to observables and nydobservables as if they both deserve the same factual status. For example, Norton talks about "the observed

[99] Van Fraassen 1980, p. 17

distribution of energy in black body radiation", and further treats this as fact.[100] However if we consider DONUD, it is a priori unjustified to count a statement about the nydobservable as *fact*. Whether or not such statements *are* facts is something to be established in this debate! We should avoid talking about nydobservable statements as if they are facts unless we have an acceptable way of establishing facts not bothered by the problems put forth in my thesis. But we *do* have at least one such a way of establishing facts: Our own senses, our own experience, can play that role well for reasons mentioned earlier. We just need to be aware that, until shown otherwise, we can only talk about facts if they are about the observable world. For instance, by our experience alone we can determine it is a fact that this sentence begins with "for instance". And that reading this sentence a week later does not matter for this fact. And to refer to Norton's statement, we can regard as fact the observable outcome of some experiment that is understood to determine the distribution of energy in black body radiation, for instance that a certain graph has a certain shape. But we cannot a priori connect with that the claim that we have then observed the distribution of energy. That said, it can still be the case that these observable facts are a sound basis for the notion of empirical support of some nydobservables, and that those nydobservables might in this way reach a fact-like status.

So when Laudan and Leplin make the point that different empirically equivalent theories are nonetheless supported in different degrees by the same set of scientific facts, this at first seems like a legitimate concern for the antirealist.[101] If there is a well developed concept like empirical *support* it may provide a solution to the underdetermination problem by way of *reasonability* or *rationality*, just as Laudan suggested. I'll repeat that support does require more than pure deduction. In my explanation of induction it even requires more than induction: for observable (A's which are B's) only suggest by way of induction universalities *amongst themselves* –that (*all* A's are B's)-, but *not* that there is a principle C from which it follows that (all A's are B's). We now need to look at a few things: 1) does the concept of support provide reason enough to solve the underdetermination problem in a way which allows us to infer realism? 2) Does my narrowing down of the concept of fact allow the same kind of support to function as Laudan and Leplin suggest?

To answer question 1 we need to have some idea what is meant when the concept of "support" is invoked. Intuitively plausible is the idea that the observation of an instance of some generalized statement provides support for the generalized statement. The statement (this observed A is a B) is an instance of the generalized statement (all A's are a B). We can broaden this idea of a

[100] Norton 1993, p. 4
[101] Laudan and Leplin 1991, p. 450

consequence by looking at all statements which are deducible from a principle C. Now we can say that the observation of an instance which is a consequence of a principle C provides support for C. For instance, the appearance of the null-result of the Michelson-Morley experiment provides support for Lorentz' contraction hypothesis, because the null-result is a consequence of Lorentzian mechanics of the ether and the contraction hypothesis. This is nothing but the standard scheme of the hypothetico-deductive method, in which you look at logical consequences of a theory or hypothesis. If you confirm these consequences in experiment, it lends support to the theory or hypothesis which entails these consequences. This system of support is haunted by all the problems we have sketched above regarding underdetermination, since a logical consequence of a hypothesis can possibly be derived from other hypotheses. After all, the null-results of the Michelson-Morley experiment also supported the hypothesis that the velocity of light has a constant value, regardless from which frame of reference it is determined, contrary to the hypothesis of the ether in which light has a certain velocity only in the rest frame.

What Leplin and Laudan want to show is that support typically implies more than what I sketch here. They attempt to show with some rather straightforward examples that there are supporting statements which are not consequences, and that there are consequences of some statement which do not support that statement. This would show that the *class* of supporting instances is not necessarily the same as the *class* of empirical consequences.

The question for purposes of the realism debate is whether or not this form of support can track truth value. Because of DONUD, we cannot use the empirical consequences of a theory alone to track truth on the nydobservable level; that is, empirical consequences alone do not provide enough justification for believing in the truth of a statement about the nydobservable. This leaves us in the dark about the truth value of nydobservables if we are left with nothing but checking empirical consequences of a theory. However, if the concept of support allows us to track truth we have a way out of the whole mess. While Laudan and Leplin attempt to show how empirical consequences and supporting instances cannot be a priori identified with each other, they do not provide any formalized notion of support. They do have a plethora of examples of supposed supporting relations which do not fall under the category of empirical consequences. I do not want to argue that because it lacks a formal notion of support it is a weak account, as there may be such a thing even without us being able to formulate exactly what it is or support might work differently in different situations which would make it valid but beyond formal notation. In the absence of a formal notion, we should examine their examples of valid inferences to see to what degree they lend credence to the idea that support tracks truth.

The first example is that of a "hackneyed case of black crows. [...] Previous sightings of black crows support the hypothesis that the next crow to be sighted will be black, although that hypothesis implies nothing about other crows."[102] Is this an example of support which tracks truth? If we accept induction as truth-tracking inference it certainly seems to be a valid case of truth-tracking support. This is not surprising, I think, as it is explicitly a case about observables. Indeed, this is not an example of support for or from nydobservables which tracks truth. On my antirealist reading I would infer, based on experience, that induction seems to be a valid form of reasoning here, as long as we experience the relative success of our use of induction here. In other words, we learn from experience that induction seems to be successful about crows being black. This does not challenge antirealism. Regardless, this example seems a mere trick; by restricting the hypothesis to 'the next crow' we avoid the logical consequence that other crows are black, but the real hypothesis we want to consider, and which is what the debate is about in the first place, is the hypothesis about *all* crows being black. The previous sightings of black crows support that the next crow to be sighted will be black because they support the hypothesis that all crows are black.

Next, they show a number of historical cases of supporting instances which were on their reading not empirical consequences of the hypothesis supported by these instances. I will cite the passage here in full for clarity:

"Consider, for instance, the theory of continental drift. It holds that every region of the earth's surface has occupied both latitudes and longitudes significantly different from those it now occupies. It is thereby committed to two general hypotheses:

H1: There has been significant climatic variation throughout the earth, the current climate of all regions differing from their climates in former times.

H2: 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 rocks from earlier periods.

During the 1950s and 1960s, impressive evidence from studies of remnant magnetism accumulated for H2. Clearly, those data support H1 as well, despite the fact that they are not consequences of H1. Rather, by supporting H2 they confirm the general drift theory, and thereby its consequence H1.

Similar examples are readily adduced. Brownian motion supported the atomic theory-indeed, it was generally taken to demonstrate the existence of atoms - by being shown to support statistical mechanics. But, of course, Brownian motion is no consequence of atomic theory. The increase of mass with velocity, when achieved technologically in the 1920s, supported kinematic laws of relativity

[102] Laudan and Leplin 1991, p. 461

which to that point continued to be regarded with great suspicion. J. J. Thompson's cathode ray experiments in the 1890s were important evidence for a host of theoretical hypotheses about electricity that depended on Lorentz's electro-atomism. Phenomena of heat radiation were used by J. C. Maxwell in the 1870s to support the kinetic molecular theory, which did not address the transmission of heat energy across the space intervening between bodies. The emergence in the 1920s of evidence showing heritable variation supported Darwin's hypothesis about the antiquity of the earth, although that hypothesis entailed nothing about biological variation. Contemporary observational astronomy is replete with indirect methods of calculating stellar distances, whereby general hypotheses in cosmology acquire support from facts they do not imply about internal compositions of stars."[103]

Another example is perhaps most telling:

"[…] a person hypothesizes that coffee is effective as a remedy for the common cold, having been convinced by finding that colds dissipate after several days of drinking coffee. The point here is that the very idea of experimental controls arises only because we recognize independently that empirical consequences need not be evidential; we recognize independently the need for additional conditions on evidence."[104]

The idea is here that the found evidence $e \coloneqq$ "finding that colds dissipate after several days of drinking coffee" does not support the hypothesis $H \coloneqq$ "coffee is effective as a remedy for the common cold" even though e is an empirical consequence of H. I think this and the historical examples above show a degree of miscommunication about the relationship between empirical consequence and hypothesis; in other words, they use the word 'support' in a way which does not mesh with how it is commonly understood. Imagine a world in which we know nothing about the common cold and coffee except that the former makes us feel bad and the latter is a black liquid which is a common drink. We find that whenever we drink the mysterious black liquid we get better from the cold. Does that not give some support to H? I think it is only because we already have the knowledge that common colds usually dissipate anyway, regardless of drinking coffee, one would say that e does not support H. But even in the light of this knowledge I would hold that e does support H, but that repeated findings of e do not make us believe H because we have different reasons to dismiss H in the first place.

A Bayesian analysis is helpful in this situation. According to a Bayesian analysis taking $P(H|e) = P(H)* [P(e|H)/P(e)]$ finding the instance e would set $P(e|H)/P(e) > 1$ since $P(e|H) = 1$ and $P(e) < 1$ if we assume that it is not a certainty that after drinking coffee the cold dissipates; after all a person can die

[103] Laudan and Leplin, p. 462
[104] Laudan and Leplin, p. 466

of the common cold. This would result in e supporting H since $P(H|e) \geq P(H)$[105] but since we don't believe H for other reasons we have $P(H) = 0$ and as a result $P(H|e) = 0$. We could make this analysis even more complicated by assuming that we have an open mind towards H so $P(H)$ is really small but not zero. This would have to involve comparison with other more accepted theories, so let's call all our knowledge about the subject the background knowledge B. Then we have $P(H|e\&B) = P(H|B)^* [P(e|H\&B)/P(e|B)]$. Since $P(e|H\&B) = 1$, as the background knowledge changes nothing about e being an empirical consequence of H, and $P(e|B) < 1$, as given the background knowledge a person might still die of the common cold, we have $P(e|H\&B)/P(e|B) > 1$ so that $P(H|e\&B) > P(H|B)$ iff $P(H|B) > 0$[106]. However we know that $P(e|B)$ is very close to 1 because of the background knowledge, and as a result $P(e|H\&B)/P(e|B)$ is very close to 1. The result is that e does support H but only *by a very small degree*. We have other hypotheses which are more *strongly* supported by different evidence.

What is made clear from this example and the historical cases is that this concept of support involves *more* than just the facts. When Leplin and Laudan refer to support it is always in a wider context in which in some way the evidence seems to support a hypothesis and other hypotheses or evidence has already been dismissed. They point out this fact because they want to show how failure to accept a notion of support in a wider context has led to a sterile way of approaching epistemic relations by looking at semantic relations only, such as 'theoretical' versus 'observational' and support relations by way of confirming empirical consequences. But this is my point of contention: if their concept of support involves *more* than the facts, the question whether or not certain evidence *supports* a given hypothesis requires further discussion on justification; namely, justifying whatever we use outside of the facts.

As an aside, some of the historical examples come across as misunderstood. Consider the example of Brownian motion; it is generally thought among physicists that Brownian motion *is* a consequence of atomic theory. Or the experimental discovery of the increase of mass with velocity; despite the suspicion towards kinematic laws of relativity the increase of mass with velocity is still a *consequence* of those laws. I wonder what suspicion towards a theory has to do with support at all.

On my reading Leplin and Laudan seem to perform a sleight of hand. While shifting the focus of attention to the a posteriori judgments of evidence they hold that this solves the problem of underdetermination. While introducing the valuable notion of support, they show in no way how this idea of support effectively solves the underdetermination problem. The example of the coffee remedy is so striking because it illustrates exactly this: their support comes

[105] Equal iff $P(H) = 0$.
[106] If we have an open mind $P(H|B) > 0$ is given.

about after all epistemological steps have been taken. They do not discuss whether or not these steps were justified, or at the very least track truth or not. I think any meaningful discussion of the concept of support should at the very least include a discussion whether or not a particular relation of support tracks truth, or that we have sufficient reason to believe that it does. Furthermore, I think their concept of support is misleading. In the coffee remedy example *e* does not suddenly 'stop' supporting *H* because we find it is insignificant and dismiss it as unimportant a posteriori. The evidence supports H, but to such a low degree we look at other hypotheses instead, and this is after we *learned from experience* that the cold generally dissipates in a few days.

In the coffee remedy case the solution to the underdetermination problem seems straightforward, because it is a case about the observable world; the answer lies in our experience. Coffee is for all purposes not a remedy for the common cold because we seem to heal just fine without it. The point of this thesis is that this straightforwardness is lacking when discussing the nydobservable world. If support is going to provide a supplement for experience it will have to bring a bit more to the table than what Leplin and Laudan provide in their article. Support comes either from confirming empirical consequences of a hypothesis, contrary to what Leplin and Laudan claim, or support does involve more than merely looking at the facts. In the former case Leplin and Laudan simply misrepresent the notion of support, in the latter case the discussion should be about what else is involved. Though I think Leplin and Laudan do partly misrepresent support, most notably in the common cold example, they provide enough reason to investigate the second option. This will be undertaken in a further section, but first I will take a look at Bayesian inference, since it provides a specific form of support.

Bayesian attempts at cracking the DONUD code

A fully fleshed out theory of support is that of Bayesianism. Bayesianism aims to use probability theory to give a precise account of how inductive support works; the *probability* of a hypothesis can be 'updated' relative to incoming evidence. Bayesianism shows precisely how this process of updating works; the mechanics of Bayesian support are given mathematically. The central idea of DONUD is still a problem for Bayesianism; evidence updates various different hypotheses in the same way as long as they all entail the same evidence. The question is: Given a precise notion of support, does Bayesianism have the tools to solve DONUD? In addition, the interpretation of the *probability of a hypothesis* plays a role. The prevailing interpretation is a subjective personalist interpretation; the probability of a hypothesis formalizes the degree of belief a person has in said hypothesis.

Earman gives an approach based on an attempt to objectify Bayesian posterior probabilities by use of convergence to certainty, or merger of opinion results, despite the subjective interpretation, and investigates its results.[107] Merger of opinion is one of the more popular Bayesian moves to go from a subjective personalist interpretation towards objective knowledge. The basic idea is that even though you begin with highly subjective degrees of belief in a hypothesis, by updating your beliefs in a rational way, namely using Bayesian procedures, the posterior probabilities or the belief in that of hypotheses after incoming evidence will converge and the subjective prior probabilities will matter less as more evidence is incoming. This is called the *swamping* of the priors. So in the long run, given certain incoming evidence, we will all have the same degree of belief in a given hypothesis, of course within a small margin which will diminish as more evidence comes in. The goal is to move to a significantly large amount of evidence -which approaches the effects of an infinite amount of evidence- so that all our degrees of belief in the real world will be equal up to an insignificant margin. After the convergence, knowledge is objective in the sense that everyone agrees, ignoring insignificant differences in degrees of belief. So we could say that in this way Bayesian procedures *objectify* degrees of belief. Note that this is not objective knowledge in the sense of being necessarily about the reality of the world, but it is objective in the sense that there is agreement among everyone, as long as they follow Bayesian rules of inference.

This would hypothetically solve the problem of underdetermination in the sense that given enough evidence, one of the available hypotheses or theories about a single domain would be selected among all the others. Eventually, given enough incoming evidence and Bayesian updating protocols, we would all agree on a *single* hypothesis of that domain out of many possible ones. The problem with this convergence is that results are only guaranteed in the long run. Usually, science experiences significant changes regarding currently accepted theories before anything justifiably called “the long run” is reached. But beyond this practical matter, it is not even clear *how* all persons will come to agree. It is technically possible for scientist X to believe that hypothesis H nears certainty after 10 years of incoming evidence, while scientist Y will only reach this degree of belief after 10000 years. The convergence only guarantees that for a given hypothesis H, a given pair of persons with certain prior degrees of belief in H and an $\epsilon > 0$ there is an N so that after seeing N pieces of evidence two people will agree on the probability of H to a maximum difference of ϵ. There can still be an extremely large variation in how large the N needs to be for different ϵ, different pairs of persons or different H and this can be understood to mean a large

[107] Earman 1992, chapter 6.

degree of subjectivity is involved. What Earman wants to find is a convergence which is stronger than this, so it limits the amount of subjectivity involved.

Earman introduces the *Gaifman and Snir Theorem*, which is about statements in a language $\mathcal{L}$ which consists of empirical predicates and function symbols combined with first-order arithmetic.[108] He looks at $Mod_{\mathcal{L}}$, the class of models of $\mathcal{L}$, which are possible worlds described by $\mathcal{L}$. To put this in terms of everyday experience; an example of a model of $\mathcal{L}$ Earman gives is "if '*P*' is an atomic empirical predicate, '*Pi*' might be taken to assert that the *i*th flip in a coin flipping experiment is heads."[109] We take a class of sentences $\Phi \subseteq \mathcal{L}$ which *separates* $Mod_{\mathcal{L}}$, that is, for any two distinct w_1, $w_2 \in Mod_{\mathcal{L}}$ $\exists$ $\phi \in \Phi$ such that $w_1 \in mod(\phi)$ and $w_2 \in mod(\neg\phi)$. Here $mod(\phi) \coloneqq \{w \in Mod_{\mathcal{L}}: \phi \text{ is true in } w\}$ and let ϕ^w be ϕ or $\neg\phi$ according as $w \in mod(\phi)$ or $w \in mod(\neg\phi)$. If all possible worlds described by $\mathcal{L}$ are indeed separated by some set Φ that means that there is no possible world described by $\mathcal{L}$ that agrees with another possible world about all the sentences $\phi \in \Phi$. Thus, Φ becomes a set of evidence by which we can determine in which of these possible worlds we happen to live. Imagine a set of a maximum of 2^N worlds in which, among other things, a coin is flipped N times. Then the set of evidence Φ which records the N flips of the coin separates the set of worlds exactly if each world has a unique outcome of the N flips, since for each pair of worlds there is at least one flip which was heads in the one world and tails in the other.

We obtain a probability space ($Mod_{\mathcal{L}}$, ,) with the set of *events* generated by $\{mod(\phi) \mid \phi \in \mathcal{L}\}$, and the corresponding probability function satisfying the conditions of being a probability function, such as countable additivity. An event in could, for instance, be all the imaginable worlds described by $\mathcal{L}$ in which the *i*th flip of a coin is heads. The idea here is to give a proof of a theorem which is valid for all the models of $\mathcal{L}$, so the theorem is valid for all imaginable worlds described by $\mathcal{L}$. In particular, it would be valid for the world which we happen to live in, given that ours is describable by $\mathcal{L}$.

Now we set $Pr(\phi) \coloneqq$ $(mod(\phi))$, $Pr'(\phi) \coloneqq$ $'(mod(\phi))$, with Pr the typical degree of belief of a given person and Pr' the typical degree of belief of another person, with the remark that these are typically not the same. Notably, in Bayesian procedures such $Pr(\phi)$ depend on the prior probabilities, and typically a pair of persons does not start out with the exact same degrees of belief in a statement. Now Gaifman and Snir theorem says for any sentence ψ of $\mathcal{L}$:

1. $Pr(\psi / \&_{i \leq n} \phi_i^w) \to 0$ or 1 (according as ψ is true or false in w) as $n \to \infty$ for all ψ such that $Pr(\psi) \neq 0$.

[108] Earman 1992, p. 146
[109] If P means "is heads", see Earman 1992, p. 145.

2. if Pr and Pr’ are equally dogmatic, that is (A) = 0 iff ’(A) = 0 for any A $\in$, then $\sup_{\psi}[Pr'(\psi/\&_{i\leq n}\ \phi_i^{w}) - Pr(\psi/\&_{i\leq n}\ \phi_i^{w})] \rightarrow 0$ as $n \rightarrow \infty$ for all ψ such that $Pr(\psi)\neq 0$.

Let’s go through the technical language and see what this means for empirical theories in a personalist Bayesian interpretation. Firstly, we presume that Φ is a set of evidence[110], summing up the *available* evidence in the form of observation sentences ϕ, and that any $Pr(\psi)$ is a measure of personal belief in statement ψ. Usually we would call ψ a hypothesis *H*. If Φ separates $Mod_{\mathcal{L}}$ then we can conclude, using (1), that our belief in a given ψ tends to head to zero or one as it is true or not as we update our beliefs given incoming evidence in a Bayesian way. This is roughly equivalent to the claim of merger of opinions, with the problems noted earlier.

Adding (2), it states that not only do two different people each come to a belief in ψ as true or false, they will also come to agree more about an *entire class* of possible hypotheses and more as they condition to more incoming evidence. So this part of the theorem prevents the situation that two people agree on some hypotheses while still having wildly differing opinions about other hypotheses. Based on (1) it was still possible to come to agreement about H_1 after 10 years but come to agreement about H_2 after 10000 years. Part 2 limits such disagreement to the extent that there is an H_i such that if N is the amount of evidence needed to come to agreement about H_i within degree $\epsilon > 0$ then for *any* H_j there will be agreement about H_j within degree $\epsilon > 0$ for N. However, I think this part of the theorem is not of any special interest. There is no guarantee that for H_j the N required isn’t extraordinarily large, which does not remove the initial problem. Even if two different people eventually come to agree about a whole class of sentences, it might still be the case that these people come to agreement about H_1 after 10 years but come to agreement about H_2 after 10000 years.

There are a few other things to note: the theorem does not state that each *different* pair of persons will come to agree in the same way, even if they come to agree at some point, and this also leaves room for some degree of subjectivity. For instance, we could have all the evolutionary biologists agreeing about the validity of evolution theory after 200 years of incoming evidence, but more sceptical organic chemists thinking differently. Earman admits as much in a footnote.[111] More importantly, the theorem requires that Φ separate $Mod_{\mathcal{L}}$, which in less technical language means that for every possible pair of hypotheses there is an empirical statement which is true for the one and false for the other. This is exactly a point of attack for an advocate of underdetermination; the problem of underdetermination is that there is *not* a set of evidence available which will separate *all* our hypotheses like this, because hypotheses typically have more to

[110] Which Earman calls an ‘evidence matrix’.
[111] See Earman 1992, p. 147 note 13.

say than statements about observable events. In other words, the theorem is limited by its very nature to statements ψ which contain no purely theoretical terms.

Earman realises this and tries to extend the language $\mathcal{L}$ to contain theoretical terms.[112] Once theoretical terms are added Φ will not serve to separate $Mod_{\mathcal{L}}$ since it is possible to have empirical equivalence with differing theoretical statements. Now Earman calls incompatible theories T_1 and T_2 *weakly observationally distinguishable* (wod) for models MOD just in case for any w_1, $w_2 \in$ MOD, such that $w_1 \in mod(T_1)$ and $w_2 \in mod(T_2)$, there exists an observation sentence O such that $w_1 \in mod(O)$ and $w_2 \in mod(\neg O)$. If $\{T_i\}$ is a partition of theories that are pairwise *wod* for MOD = $Mod_{\mathcal{L}}$, then for any $T_j \in \{T_i\}$ (1) is valid so $Pr(T_j/\&_{i\leq n}\ \phi_i^w) \rightarrow 0$ or 1 for a suitable $\Phi=\{\phi_i\}$ as long as T_j isn't rejected right off the bat.

But, as Earman mentions, failure of *wod* could be taken as precisely what underdetermination by observational evidence means. What I would rather state is that underdetermination *of any respectable form* implies failure of *wod*. If there would be sentences $\{O_i\}$ such that each $O_j \in \{O_i\}$ would be an observable difference between any two hypotheses, then there would be no underdetermination because by using $\{O_j\}$ we can drop all the hypotheses which have false observational results. *This is exactly what* $Pr(T_j/\&_{i\leq n}\ \phi_i^w) \rightarrow 0$ or 1 *means*!

So we see that despite attempts to objectify Bayesian inference, Bayesianism does not solve the root problem. This isn't surprising. Consider the basic Bayesian scheme:

$$P(H|E) = P(H) * P(E|H) / P(E)$$

Notice that this isn't very different from hypothetico-deductivism. As you find the evidence E, it increases the probability of hypothesis H as long as H entails E. Bayesian inference adds a quantitative measure of determining how much the probability of H increases. You can increase the content of the Bayesian scheme by considering other hypotheses or background knowledge B:

$$P(H|E\&B) = P(H|B) * P(E|B\&H) / P(E|B)$$

When considering such Bayesian schemes, whether or not E provides positive confirmation to H depends on the ratio $P(E|B\&H) / P(E|B)$. But that ratio is fully dependent on the factors involved. *This is the case for every Bayesian comparison.* Now it is either the case that B is relevant for the appraisal of E or it is not relevant. In the first case the underdetermination problem is never solved. In the second case the issue is to determine the terms. For this, I claim, one needs to resort to scientific values.

Earman gives the following example by Dorling:

[112] See Earman 1992, p. 151 as well as note 19.

"Suppose that theory T consists of core hypotheses T_1 and auxiliary assumptions T_2; that T_1 & $T_2 \vDash E'$; and finally that nature pronounces E, which is incompatible with E'. Dorling assumes that T_1 is probabilistically irrelevant to T_2 (that is, $P(T_2|T_1) = P(T_2)$), that the priors $P(T_1) = k_1$, and $P(T_2) = k_2$, satisfy $k_1 > k_2$ and $k_1 > .5$, while the likelihoods $P(E|\neg T_1$ & $T_2) = k_3$, $P(E|T_1$ & $\neg T_2) = k_4$, and $P((E|\neg T_1$ & $\neg T_2) = k_5$, satisfy $k_3 \ll k_4$, $k_5 \ll 1$. Then Bayes's theorem shows that the blame falls more heavily on the auxiliaries T_2 than on the core T_1." [113]

However, and Earman points this out right away, this requires a particular assessment of all the terms involved. The values k_1 to k_5 are not determined by empirical considerations, and I think one requires the use of scientific values to determine them. Earman thinks that '[i]n general, none of the factors involved has an objective character, and a large variability can be expected in the values assigned by different persons.'[114] As he notes, the values may be assigned differently especially if we consider possible future rivals in the vein of Stanford.

Thus, a solution to this problem requires objectifying Bayesian probabilities. But as noted above all attempts provided by Earman fail to avoid the underdetermination problem. I think that is because Bayesian updating is merely a glorified version of hypothetico-deductivism, and while it adds a valuable quantitative aspect, it is wholly neutral on the issue of scientific values and appraisal of the actual probabilities. Usually, the crucial term is the likelihood of the evidence P(E|B) or the prior probability of the hypothesis P(H|B). Both values are difficult if not impossible to determine. For instance, an often cited reason for believing H given E is that P(E|¬H) is very small.[115] But if we allow Stanford's scepticism as noted above, we cannot determine that P(E|¬H) is small. I would in fact argue that P(E|¬H) is impossible to determine without having access to an exhaustive set H' which together make up the set ¬H.

Values in adductive inference

We have seen that an account of support or confirmation is either just as powerless as deduction or hypothetico-deductivism to avoid DONUD, or relies on something more than just the bare relations of fact and theory. In this section I aim to discuss what critics have in mind when they propose various adductive inferences. This section could be seen as an explication of what Duhem's *bon sense* possibly entails. Like Laudan, Leplin, Collins et alii, Duhem himself suggests that the ultimate decisions in science are made by something outside purely logical considerations.

---

[113] Earman 1992, p. 83
[114] Earman 1992, p. 84
[115] For instance, Norton argues this point, see below.

T.S. Kuhn proposed in 1973 that there are certain scientific *values* which influence theory choice, though the sentiment can already be found in his revolutionary book The Structure of Scientific Revolutions published some 10 years earlier. He suggests that scientific values should not be seen so much 'as rules, which *determine* choice, but as values, which *influence* it.'[116] Kuhn suggests five values which are the most important: accuracy, consistency, scope, simplicity and fruitfulness. Of these, I'd suggest that only the latter three should be seen as examples of adductive inference, as the other two are aimed at empirical considerations.

That such values should play a role in scientific decision making makes sense. Kuhn has explained well how a new paradigm can seem empirically weak compared to an older paradigm. Nevertheless, proponents keep working on their new ideas, and eventually the new can become empirically better than the old. In the mean time non-empirical arguments are used to persuade opponents. However, critics of Kuhn have argued that this means Kuhn does not allow science to have rational theory choice.[117] I will not discuss the point of rationality but I will say this: Regardless of whether or not scientific theory choice is always based on empirical or rational argumentation, it seems evident that what Kuhn calls values sometimes play a role. I will understand 'value' to mean any property of a statement, theory or hypothesis which is not empirical but is used to give epistemic judgment about the statement, theory or hypothesis. Promoting a hypothesis by way of praising it for having a value of this kind is then a form of adductive inference.

In fact, realists often use the strategy of involving values to argue for realism. For instance, J.D. Norton is a realist who explicitly uses this strategy. Norton makes two central claims in his argument. Firstly he claims that an account of underdetermination can only function based on an 'impoverished account of induction'.[118] Secondly he claims that an argument based on empirical equivalent theories is self-defeating.

Immediately I want to point out that Norton is guilty to the same degree as Leplin and Laudan of giving an account of underdetermination which is in my eyes too strong. Norton says the following:

'…underdetermination is explicated as the assured possibility of rival theories that are at least as well confirmed as the original theory by all possible data or evidence.'

And later on the same page:

[116] Kuhn 1998, p. 111, emphasis mine.
[117] See Kuhn 1998, p. 102-103
[118] Norton 2008, p.17

'The underdetermination thesis is much stronger; it asserts that all theories are beset with [the problem that part of a theory transcends evidential determination.]'[119]

Firstly, DONUD does not entail an availability of *well-confirmed* rivals. It *leaves open* the possibility of equally well confirmed rivals, to be sure, but it does not assert that this *will* always be the case. Secondly, what it means to be 'at least as well confirmed' is at this point of the debate a vague notion which does not necessarily contain any truth-tracking power. Claims to the contrary still need to be made plausible. Thirdly, it only claims that the second citation is true *in principle*, as a natural result of the relation between theory and experiment. I would also argue that it is *trivially* true that there are parts of any theory that transcends evidential determination; even the simplest theory rests on assumptions which defy evidence. For example, special relativity *assumes* that the velocity of light in a round-way trip determination of simultaneity is equal regardless of the direction of the light wave. Beyond specific examples like these which I am sure can be found in any theory, there are more general assumptions like the presence of causality or the need for replication of experiments. There is no *guarantee* that scientists will actually run into the underdetermination problem in practice. I think, however, that there is ample evidence to suggest that scientists *do* run into that problem on a day to day basis. This is the reason why, I assert, in experimentation the road to a publication is difficult and fraught with problem solving. One could take for instance the process of copying the Jumbo laser in Collins' book, and I am sure experimental scientists will find that account recognizable.

In addition, I'm not sure Norton understands that the original Duhemian argument comes from the relation between theory and experiment. He refers to a logical distance between theory and experience, while to me it seems that the problem arises from experience and theory being intertwined in experimentation. That he does not think the same is evidenced in some of his examples. For instance one of the generalized claims he makes is: "Many accounts of confirmation allow evidence to bear directly on individual hypotheses within a theory than merely supporting the entire theory holistically."[120] He uses an example of observing spectral lines of helium from sunlight. This is on his reading strong evidence for the presence of helium in the sun, despite that it 'requires numerous additional assumptions about the optics of cameras and spectrographs.' I assert that Norton is begging the question here. On the one hand the determination of helium in the sun requires numerous assumptions, but on the other it is claimed we can dismiss them as bearing, at least significantly, on the confirmation relation. He does not state why we should

[119] Norton 2008, p.20
[120] Norton 2008, p. 29

dismiss them as bearing on the confirmation relation. I will repeat the Duhemian argument: precisely because we require numerous assumptions about the optics of cameras and anything else involved we cannot simply assume that the observation of what we call 'spectral lines' of helium is evidence for the presence of helium; we can only assume that given the assumptions about optics of cameras et cetera the observation of 'spectral lines' of helium is evidence for the presence of helium. Any account of support or inductive inference to the contrary would have to be made plausible in its own right. A counterexample has been given in the discussion about Bayesian inference above; it shows that while Bayesian inference at first glance allows evidence to bear on a hypothesis in isolation, further analysis shows that this requires an evaluation of the evidence which does rely on additional assumptions.

Norton argues that there are roughly three avenues of inductive inference which might be able to solve the underdetermination problem. The third family of inductive inference is that of probabilistic accounts. I have discussed Bayesianism, as the prime candidate for such accounts, in the previous section. The first family, inductive generalization, is a basic family of principles in which an instance confirms its generalization; think of enumerative induction as the prime example. Norton comments on Laudan and Leplin's argument that there is evidence for certain hypotheses which are not consequences of said hypotheses. He notes that "Laudan and Leplin proceed to display many more examples of cases of evidence that are not consequences."[121] I would like to point to my discussion of Laudan and Leplin above and simply restate that all their presumed examples are on closer inspection not good examples for this claim after all. In fact a closer inspection supports the age-old idea that evidence for a hypothesis consists of its consequences. Further, Norton mention's Glymour's bootstrap theory. It's strength is that "we have already strong, independent evidence for the needed auxiliary hypotheses, so that the inference from evidence to hypothesis can be made deductively without taking any further inductive risk."[122] However, if we consider DONUD, there is no basis for assuming that the evidence is independent, since any evidence for such auxiliary hypotheses is itself subject to the same criticism DONUD raises against the core hypotheses to be examined. This is what Collins has pointed out in his discussion on the experimenter's regress.

The second family of inductive inference that could stem the tide of underdetermination is hypothetical induction; think of the 'impoverished account of hypothetico-deductive confirmation that lies behind the underdetermination thesis'.[123] According to Norton these accounts do avoid

[121] Norton 2008, p. 30
[122] Norton 2008, p. 30
[123] Norton 2008, p. 31

underdetermination because they always have additional constraints which bear on inference. If we take a look at what kind of additional constraints Norton has in mind, we quickly find that all of them would require additional support in order to convince us these additional constraints allow us to track truth in the hypothetico-deductive model. His constraints are as follows:

1. For evidence E to confirm H, H must not only entail E but it must also be shown that if H were false E would very likely not have obtained.
2. The hypothesis being confirmed is simple, with complicated rivals thereby precluded.
3. The hypothesis supported by the evidence must also be the best explanation of the evidence.
4. The hypothesis confirmed must in addition be produced by a method known to be reliable.

These are all relatively easy to debunk but each requires some serious discussion.

Point 1 does not allow us to track truth because there is no reliable way to determine that E would not likely have obtained if H were false.[124] This additional constraint would presume that we have access to all possible H' $\subset \neg$H which could be considered rivals to H, and that we have determined that none of these are likely to determine E. In reality I'd say we typically known only a few serious rivals. The criticism to 1 stems from a simple question: "Shall we ever dare to assert that no other hypothesis is imaginable?"[125]

Coincidentally, this same argument can be used to debunk point 3 with slight modification: "Shall we ever dare to assert that no better explanation is possible?"

Whether or not point 2 allows us to track truth is an issue that requires additional discussion. Simplicity is a scientific value often discussed concerning theory choice and empirical equivalence. Lorentz vs. Einstein and Copernicus vs. Ptolemeus are two well-known examples of theory choice in which a system's supposed simplicity was used as a reason for accepting it over the alternative. Many discussions have subsequently been about what this simplicity entails and whether or not it should be sufficient reason for believing one theory over the other, if it has greater simplicity. I will say this about simplicity: A simpler theory ought to be rejected in favour of a more complicated theory if the more complicated theory reproduces observable phenomena and a simpler does not. This does not mean that simplicity cannot play a role in justification of belief in

[124] Let us set aside for now the issue that we would want E to *surely not* have obtained if H were false, instead of 'very likely not'. I am granted the disposition that this certainty is not required, because my concern is whether or not we have *good reasons, not undeniable proof*, for being antirealists.

[125] Duhem 1982, p. 190

nydobservables, but it does mean that it is not a priori clear *when* it should play a role and when not.

In response to 4, I will say that there is, at this point in the discussion, no reliable method outside of learning from experience that would fit the bill of such a constraint. Norton's aim of adding this constraint seems to be the dismissal of so called ad hoc hypotheses, since these are generated not by a reliable method but by 'artful contrivance'.[126] However, since there have been examples of ad hoc hypotheses which were nevertheless borne out in subsequent research and accepted by the scientific community, such as the hypothesis of the existence of Neptune, Lorentz' contraction hypothesis and the existence of neutrinos, it does not seem that we can dismiss all ad hoc hypotheses as false a priori. If science should reject them, then this shows that science is not in the business of finding true statements about the world, invalidating realism. If science does not reject them, then there are *unreliable* methods capable of generating true statements.

Furthermore there seems to be a case of circularity involved. What is reliability? If it is the repeated reproduction of empirical results only, then it states no more than that in order to track truth, a theory should agree with experiments. If it is not only that, then reliability is a scientific value, the use of which begs the question. Reliability would then also imply the repeated production of theoretical statements which are true. But whether or not they are true is the point of discussion.

Kukla defends the idea that the argument for empirical equivalents can be supported by *algorithmic construction*. The strategy for algorithmic construction is the following: Take a theory T. Now construct an empirically equivalent T' by adjusting the theoretical terms in T. The precise adjustment differs, but the key part of the algorithm is to leave the observational terms alone, thereby preserving observational equivalence. For instance, Kukla mentions this example from van Fraassen: 'given any theory T, construct the rival T' which asserts that the empirical consequences of T are true, but that T itself is false.'[127] Another example is T'', constructed by taking T_1 to be true whenever someone is observing something, and T_2 to be true whenever nobody is observing, with the proviso that T_2 is incompatible with T_1. Even though my position based on DONUD does not presuppose the ability to construct these kind of algorithmic rivals, the discussion touches upon epistemological values and so is interesting to follow.

The point of interest is the argument against such algorithmic constructions. Kukla reports that according to critics of this strategy, rivals constructed in such a way can be rejected a priori because they fail to meet certain criteria which allow them to be proper rivals in scientific endeavour. Theories which fail to

[126] Norton 2008, p. 31
[127] Kukla 1996, p. 138

meet scientific criteria are called *quasitheories*. Then a strategy is to disallow quasitheories to play an actual role in the underdetermination argument.[128] Kukla's response to this argument is *theory-shmeory*: Call anything which is either a theory or a quasitheory a shmeory. Then we can repeat the empirical equivalence argument replacing 'theory' with 'shmeory'. But the point of Leplin et alii is that these shmeories do not all *count as rivals*, how does Kukla refute that argument? He argues that since every shmeory is at least logically possible, the rejection of quasitheories must be based on something else than purely empirical reasons.[129] As Kukla points out, this must be based on what is *part of the debate in the first place*: whether or not criteria to determine a shmeory as worthy of the title theory are epistemically relevant. Even if a quasitheory like the constructed T' is rightly rejected in scientific discourse, there is no reason why it should not play a role in epistemic considerations.

Leplin and Laudan attempt to construct an idea of theoreticity, what counts as a theory, and what does not.[130] Despite Kukla's further appraisal of criteria for theoreticity, I think the argument given in the previous paragraph is enough to show that the problem of underdetermination, whether of the form of DONUD or Kukla's contrastive underdetermination, is prior to these considerations. Scientific criteria for theoreticity and scientific values are possible *solutions to, not arguments against*, underdetermination. Underdetermination is a *fact of life* but we need to determine whether or not it is an *insurmountable* problem. The aforementioned criteria and values may well be solutions, but need to be judged on their own.

In addition, there have been examples of empirically equivalent theories which are comparable to the pairs of T and an algorithmically generated rival T' in the history of science. Lorentz' theory of electrons is comparable to Einstein's special relativity: The former can be understood in terms of the latter, if the hypothesis of the existence of the ether is appended. Kinetic theory emerged as an alternative to thermodynamics but the kinetic theory tried to model the microscopic world. Tycho Brahe's cosmological view in which the Sun revolves around the Earth was observationally indistinguishable from Copernicus' heliocentric view. Certainly as these theories developed in time, empirical equivalence ceased to be a factor as the domain of one of the theories increased, but that is not the point; the point is that it is indeed possible for theories to emerge which are both valid from a scientific point of view, thus satisfying all the criteria of theoreticity, but which could be understood as an example of a situation of algorithmic construction. Thus, it is always possible that a situation of empirical equivalence emerges, even if it is not given that such situations will

[128] This is in fact the strategy that Laudan and Leplin employ.
[129] See Kukla 1996, p. 146-147
[130] And thus what can be discarded for epistemic evaluation.

occur. I think that the mere possibility of such a situation emerging is enough ammunition for antirealists. This is what I understand Stanford argues with his New Induction.

Some of the criteria for theoreticity and scientific values that Kukla mentions are only important when generating algorithmic rivals. For instance, the parasitism criterion is characterized as: "T' is totally parasitic on the explanatory and predictive mechanisms of T … a [real] theory posits a physical structure in terms of which an independently circumscribed range of phenomena is explainable and predictable."[131] Since my argument does not rely on the generation of such rivals I will focus my attention on other criteria and values which are not dependent on algorithmic construction; these may be valid in other situations which are significant for my discussion.

Kukla discusses a criterion similar to Ockham's razor called the *superfluity criterion*. The comment by Kukla that this is more a value than a measure of theoreticity seems valid, since it either is a criterion for theoreticity and thus an argument for antirealism and the abolishment of all theoretical statements, or it is a value which only measures what is the better theory. After all, van Fraassen uses a similar argument to show that constructive empiricism is a stronger position than realism, but van Fraassen does not argue that realism is not a theory at all because of this superfluity.

The interesting part here is a peculiar example. A theory T is compared with a constructed rival T" which says that T is true whenever someone is observing, and another theory incompatible with T is true whenever nobody is observing. A superfluity argument implies that T" has superfluous parts which can be excised. But T has equally superfluous parts, since it says that T is true whenever someone is observing, and true when nobody is observing. My antirealism would however propose another T"', which says that T is true when someone is observing. It says nothing else. The world in which we observe *is* our reality, both statements that T continues to hold or not are totally irrelevant epistemologically. There is no experience correlating with the status of T when not observing; this means it is not part of our experience. And as I've said that science should be about our experience, since that is all there is to us, any statements about T when not observing are irrelevant.

Kukla introduces *NN*; the assumption that 'non-empirical virtues have no bearing on epistemic evaluation'.[132] The issue is about whether or not non-empirical virtues, criteria by which to judge a theory's worth in some way, are epistemically valid. For example, we may wonder, given two empirically equivalent theories, which theory provides us with a reason to believe it above the other, even on the unobservable level. If we exhaust the empirical resources

---

[131] Kukla 1996, p. 148

[132] Kukla 1994, p. 157, Note that what Kukla calls 'virtues' I have referred to as 'values.

towards this goal, we have to rely on something else. We might select one over the other because it has fewer controversial assumptions, or does not manage to explain all its phenomena from one equation, or that it has elements which can be excised without losing empirical prowess. Whatever reason we use, we can quickly see there are not directly empirical considerations. The question is whether or not they can, as non-empirical virtues, play a role epistemologically. For instance, can the simplicity of a theory be reason for believing it to be true, or can we believe the simpler of two empirically equivalent theories to be true?

NN is thus the hypothesis that these kind of virtues do not have epistemic value. The first thing to note that such virtues can only be fruitfully applied in a situation of theory *choice*. We can readily see that a criterion like simplicity fails when applied to a single hypothesis; judgements of simplicity can only be made gradually. A single hypothesis cannot be 'simple', it can only be 'simpler than'. Scientists are primarily guided by empirical considerations; non-empirical virtues often play a role but on the long run are subjugated to empirical considerations. For instance, the philosophical problems connected with quantum mechanics are hardly a deterrent for physicists; *as long as we can use it to calculate* is enough reason to use the theory for some. This is not to say that empirical virtues are always the primary concern; there may be cause to abandon strict empirical adequacy for pragmatic purposes. It is just that if this course of action does not lead to a long-term generation of increasing the domain of empirical knowledge, it is abandoned.

One antirealist answer to NN is that there are *no* such standards for the purposes of epistemic appraisal. Scientific values, when valid, are only valid as pragmatic concerns. But, as Kukla points out, antirealists vary in their exact opinion about values. Van Fraassen, for example, is committed to allowing at least some criterion for epistemic appraisal on the observable level.[133] Even though Van Fraassen has aimed criticism at the use of non-empirical virtues in epistemic appraisal, Kukla points out that Van Fraassen has claimed doing so would lead to incoherence.[134] It is either the case that we cannot use non-empirical virtues at all but in that case we cannot distinguish between empirically equivalent hypotheses and neither can we do so between grue-like hypotheses, says Kukla. Or it is the case that we can use non-empirical virtues on the observable level at least, but this raises the question why we shouldn't use those virtues on the unobservable level as well.

Kukla concludes that since the first options is undesirable, likely even for Van Fraassen since it would entail an undesirable form of skepticism, the second option is the preferred one. But this option does not force *either* the position of

[133] See Kukla 1994, p. 159 and p. 168. Van Fraassen himself famously admits he is sticking his neck out.

[134] Kukla 1994, p. 163

constructive empiricism, and hence anti-realism, *nor* the position of realism. It merely states that both are acceptable and the debate is undecided.

A. Fine is a famous critic of both realism and anti-realism for that same reason. Fine notes that anti-realism relies on the same kind of inferences regarding the observable world as realism relies on regarding the unobservable world. Thus the skeptical attitude is to wonder why we can use those types of inference on the observable level but not on the unobservable level. This argument is similar to those arguments which claim that anti-realism of the form I suggested will lead to radical skepticism. I have argued against this view above, but I will respond to the following citation for clarity:

"Thus, this brand of empiricism [constructive empiricism and other kinds of anti-realism] can follow the usual lattice of inferences and reasons that issues in scientific beliefs only until it reaches the border of the observable, at which point the shift is made from belief to acceptance. But the inferential network that winds back and forth across this border is in no way different from that on the observable side alone."[135]

According to DONUD and my remarks about learning from experience at the beginning of this chapter, there is a difference in the 'inferential network' on the observable level and the unobservable level. The network does not 'wind back and forth across this border' precisely because *experience* is the main difference between the two levels. That is, whenever we apply adduction we refer to values to come to a different kind of theory appraisal than we would purely on empirical grounds. The question of justification hinges on those values. I claim that justification is given on the observable level but not on the nydobservable level, precisely because we can learn from experience which value is justified for what situation on the observable level. There is no such independent justification on the nydobservable level.

We can learn from experience when we can reliably apply a scientific value. For instance, we learn from experience that we can use induction in some cases, like the sun rising in the morning, but not in others, like a neighbour rising from his bed in the morning. When we wake up and find our alarm clocks blinking, we can rely on the simplest explanation: the power went off. In another situation, such as the cause of the power outage, we cannot so easily apply this strategy. In other words, values do not have *universal* application, and we do not seem to be able to determine rules which allow us to infer when and when not to use which values. On the other hand, based on the principle of learning from experience we do seem to be able to use them, albeit tacitly. *The use of scientific values is thus a form of tacit knowledge.* This in turn explains why experimenters

[135] Fine 1986, p. 167

seem to have an intuitive grasp of their experiments and instruments, even if they are not capable of articulating what it is they understand.

I believe that this is exactly what Feyerabend tried to point out in his notorious book Against Method. It is not so much that science is a completely irrational affair in which rational decision making plays no role. It is that rational decision making does not follow universal principles:

"Neither science nor rationality are universal measures of excellence. They are particular traditions, unaware of their historical grounding."[136]

"Yet it is possible to evaluate standards of rationality and to improve them. The principles of improvements are neither above tradition nor beyond change and it is impossible to nail them down."[137]

But on the nydobservable level, there is no experience to guide this process of learning. In fact, as Stanford's New Induction points out it should be expected that we are wrong about at least a few important things in our current theories. In other words, that our use of scientific values to support our 'belief' in current theoretical statements may well turn out false. We may of course be wrong on the level of the observable, but the point is that we can achieve certainty there. It is *in principle impossible* to achieve certainty on this on the nydobservable level. This is exactly the reason why one should have an anti-realist attitude towards science.

The no-miracle argument and abduction

Two arguments for realism, which I will characterize as use of scientific values as discussed above, that are important and must be discussed before coming to the conclusion are *abduction* and the *no-miracle argument*. Even though I have already given an argument why any scientific value does not solve the DONUD problem, I will discuss these separately, but note that my general argument above does apply for these values.

Abduction is also known as 'inference to the best explanation' and can be characterized as follows: Abduction is inference that allows one to choose one hypothesis over the other by using explanatory considerations. Abduction can be best explained by an example:

"You happen to know that Tim and Harry have recently had a terrible row that ended their friendship. Now someone tells you that she just saw Tim and Harry jogging together. The best explanation for this that you can think of is that they made up. You conclude that they are friends again. […] It does not follow logically that Tim and Harry are friends again from the premises that they had a terrible row which ended their friendship and that they have just been seen jogging

[136] Feyerabend 1993, p. 214
[137] Feyerabend 1993, p. 230

together; it does not even follow, we may suppose, from all the information you have about Tim and Harry. […] What leads you to the conclusion, and what according to a considerable number of philosophers may also warrant this conclusion, is precisely the fact that Tim and Harry's being friends again would, if true, best explain the fact that they have just been seen jogging together."[138]

We use abduction to come to the conclusion that Tim and Harry are friends again. In science the argument would be similar:

'If electrons are real, this would explain the success of theories involving electrons. We have successful theories involving electrons; if electrons were real we would be able to explain the success of these theories. Thus, the existence of electrons is the best explanation for the success of our theories, and so we can infer that electrons are real.'

The no-miracle argument can be characterized as follows: Only realism can explain the fact that theories, even when they are constructed to only describe a given domain of phenomena, can give predictions about a new domain which bear out. That is, the success of science in giving novel predictions can only be explained if it is assumed that scientific theory is true; that its theoretical statements correspond to how the world really is. To find this predictive success without believing in the -approximate- truth of scientific theories would make this predictive success miraculous.

I will first turn to abduction. My argument given above can be repeated: since we all should have experienced at least a few cases in which what we thought was the best explanation didn't turn out to be true, abduction does not have universal value. For instance, imagine a little boy searching for his favourite toy. As he cannot find it he quickly blames his brother, since he knows that his brother frequently steals his toys. Later it turns out that their mother has been cleaning their toys and thus the toy was not stolen by the larcenous brother. Or an example from physics: Once it was thought that light being a wave would best explain the phenomena of light known, then it was thought that light being consistent of particles was the best explanation, now we aren't sure what to believe anymore. The point is that as far as explanations go, they are notoriously unreliable. So in what case should be use it on the nydobservable level? We may be able to learn when or when not to apply abduction on the observable level, but experience does not allow us the same on the nydobservable level. I see no other way in which we should believe that abduction allows us to track truth on the nydobservable level.

As for the miracle argument, as Fine points out, it presumes a hindsight look at science. Fine explains cogently that this is a result of our historical perspective:

---

[138] Douven 2011

"If, for example, we could examine the myriad attempts in laboratories around the world just (literally) yesterday to turn basic science to the production of a useful instrument, then, I think, we would find failure on a massive scale, and certainly not any overall success. [...] For the application of science involves an enormous amount of plain old trial and error; hence, it always entails an enormous amount of error. I think a reasonable historical picture would be to draw each success as sitting on top of a great mountain of failures."[139]

In other words, *there is no miracle to explain*. If we are given to a large amount of cases and we apply trial and error, we should expect that at some point we should arrive at success. That is, it is entirely possible that the success of novel predictions is not due to the corresponding theory being true or having true components, but due to scientists' ability to tweak and adjust in such a way that the experiments have the 'expected' results. That scientists are in fact capable of doing that is evidenced by their capability to produce 'success' with theories that are now considered patently false. Laudan has argued this point elsewhere and I consider his account sufficient.[140]

Fine suggests that there is an internal reason why abduction isn't all that it seems. Fine claims that any explanationist argument which infers the truth of theoretical statements from their instrumental reliability can be subverted for the instrumentalist cause by replacing *truth* with *reliability*. So instead of inferring that electrons are real from the instrumental success of using the term electron, we may infer that using the term 'electron' is reliable. This argument is similar to that of van Fraassen whose central argument for constructive empiricism is that it is a position which says less about the world than realism.[141] In other words, van Fraassen promotes constructive empiricism on the basis that it is logically weaker, and thus a safer position, similar to Fine's argument here.

But according to Fine the no-miracle argument is in an even worse position, namely that it is fundamentally question-begging; it assumes the truth of theoretical statements in its argumentation. To reconstruct Fine's argument, consider the following abductive inference.

'If electrons are real, this would explain the success of theories involving electrons. We have successful theories involving electrons; if electrons were real we would be able to explain the success of these theories.'

So, the realist concludes, we have a reason for accepting the reality of electrons; namely that it allows us to explain why theories involving them are successful. Now consider the same inference, but we replace 'electron' with

[139] Fine 1986, p. 153
[140] See Laudan 1998b
[141] Fine 1986, p. 154-155, see metatheorem I. Also see van Fraassen 1980, p.73.

'correspondence'. I refer to correspondence in the sense of correspondence between theoretical statements and reality; the realist sense of truth.

'If correspondence is real, this would explain the success of theories involving correspondence. We have successful theories involving correspondence; if correspondence were real we would be able to explain the success of these theories.'

So, the realist concludes, we have a reason for accepting the reality of correspondence, in other words for accepting the realist account of scientific theories; namely that it allows us to explain why theories involving it are successful. As Fine points out, anti-realists reject the general form of inference as used in the electron case. The correspondence example uses the same kind of inference, why should they accept it there?[142] Thus, the no-miracle argument is a mirage, an account based on an unacceptable form of abduction in addition to a skewed look at the history of science.

Conclusion: DONUD cannot be solved on the nydobservable level

We have seen that attacks on underdetermination by realists have taken two courses: one the one hand the claim that a more sophisticated account of support or confirmation or inductive inference allows us to solve underdetermination, on the other hand the claim that scientific values play a role in theory choice. Leaving aside the point that theory choice is not equivalent to truth tracking, I first notice that existing accounts of support, confirmation or inductive inference fail to deliver as promised, unless they also employ the use of scientific values. Leplin and Laudan give a very weak account of support which I have shown to be invalid unless one employs values, Norton explicitly refers to values and Bayesian inference remains neutral on the subject because it is inherently not much different than hypothetico-deductivism. Where Bayesian attempts do manage to provide an apparent solution to underdetermination, it is only because of appraisals of prior probabilities of hypotheses and likelihoods of evidence, which cannot be done without the use of scientific values, whether through reference to a 'background knowledge' or not.

Thus the issue becomes that of the truth tracking ability of scientific values. It is sometimes suggested that antirealists have not given any reason why realists should not use scientific values if proof is not given that scientific values cannot play a role in theory choice.[143] I argue that this is a fallacious argument; the realism issue is about *justification.* If realists cannot give justification why we should allow scientific values to play a role in determining our *belief*s about the

[142] Fine 1986, p. 160-161

[143] See for instance Kukla 1994 and Fine 1986

theoretical statements of a theory (namely that they are true statements) then *realism itself* is not a justified position. We should be anti-realists in that case.

Nevertheless I have looked at several values and concluded that none of them have universal value. Looking at Feyerabend and various historical works, it is reasonable to assume that, until proven otherwise, *no* scientific value has universal value. Indeed, one can readily think of any scientific value and think of a case in which use of it would lead to a wrong conclusion. In any case, I believe the ball remains in the realist court to prove that any scientific value has universal value.

Experience, however, allows us to learn in which cases the use of a value is warranted. This is not something I think we can explain philosophically, precisely because there is no universal way in which values are used. My brain seems perfectly capable of understanding that the sun will come up tomorrow even if I cannot give a conclusive reason why this should be so. Indeed I believe that there is nothing more than learning from experience at play here. How do we learn from experience? Suffice to say that this thesis is not the place for that discussion, but I do think Bayesian inference is a good candidate, even if it only gives a very schematic model of learning.

Experience in fact allows us to avoid the charge of inductive skepticism. There is no reason why we should doubt that the sun will come up tomorrow, because we have learnt from experience that it does. We also learn from experience cases in which induction does not work. Also, we do not need to fear radical hypotheses involving invisible elves at work everywhere which we cannot measure in any way but still produce the world as we know it. There is nothing in experience which suggests such hypotheses. To be sure, there is nothing in experience which directly suggests that molecules and electrons are real either, but that is precisely the point of this thesis; we should be antirealists and being an antirealist involves not believing to be true that molecules and electrons are real. They might exist, but we don't have enough reason to say we know that to be true.

PART III: CONCLUSION AND SOME REMARKS

Conclusion

I'll present my conclusion as a series of theses I've proposed in all of the discussion above.

- Underdetermination is mostly misrepresented in the realism debate. Realists focus on forms of underdetermination that are too strong, and also do not represent well what Duhem and Quine intended. Realists do not challenge the argumentation which is the foundation of DUD and QUD.
- A serious underdetermination account can be gained by taking Duhem's account and adding the explanation of verification that Collins has provided. This I give the name DONUD; DONUD is a serious problem for realism, and there should be an argument allowing us to solve the DONUD problem before realism can be accepted.
- Accounts that directly attack underdetermination are either flawed or constitute a straw-man argument.
- Experience is the dividing line between the observable and the nydobservable world. It provides justification for believing statements about the observable world (to be true), but not in any statements about the nydobservable world (to be true).
- Learning from experience is a sound basis for scientific investigation.
- Attempts to solve underdetermination based on the idea of support or confirmation or inductive inference fail; they are either fundamentally incapable of avoiding DONUD because they are not so different from hypothetico-deductivism, or they apply judgments of hypotheses based on scientific values. In the latter case the issue of justification hinges on justification of the values involved, but see the next point.
- Scientific values require separate justification. Realists have so far given only very weak defences of such justification. Even if antirealists have not given reason to conclude that use of values does not track truth, leaving open the *option* of realism, I argue that this debate is about *justification*. We require further justification for believing any statement about the nydobservable, and if none is given, realism should be avoided. One cannot be a realist with justification if this is not given.
- Scientific values do not have universal value; that is, they sometimes but not always track truth, whether or the observable or nydobservable level. The question is when we can rely on scientific values to track truth. I argue that we can learn from experience when scientific values track

truth and when not, and that we use this knowledge tacitly both in everyday life and in science.

- However, experience does not allow us to learn the truth about the nydobservable level. That is, since we cannot experience the truth of any statement about the nydobservable level, we have no way of knowing when a scientific value tracks truth on that level.
- DONUD is a fundamental problem on the nydobservable level, but can be avoided (but not always is avoided) on the observable level. Because of this realism is not a justified attitude in the realism debate. We should be anti-realists.
- Constructive Empiricism is not acceptable because it makes an arbitrary distinction between the observable and the unobservable in the sense that it allows scientific values to track truth on the one but not on the other; it only allows scientific values to play a role in pragmatic concerns on the unobservable level.

Some remarks about Structural Realism

During the writing of my thesis I was made passively familiar with the idea of Structural Realism. The idea is that throughout the development of science, theoretical entities have come and gone but structure has been kept intact. If that is the case, Stanford's New Induction fails regarding structure, and there remains a lot to say for the failure of DONUD regarding structure. That is, given scientific development and DONUD, there is enough reason to assume that structure is not underdetermined. The argument is similar to that of the no-miracle argument, and it avoids Laudan's devastating pessimistic meta-induction. In particular it does not defeat DONUD, but makes a central argument for it, as given by Stanford, implausible at least on the level of historical appraisal.

I will have a few remarks on it, even if this thesis does not have the aim to discuss it. First of all, I want to make clear that, as far as I have read about it, it seems to me a serious alternative to my anti-realism above. I find the central idea attractive and convincing, especially because it agrees with most of my discussion above. Structural Realism rejects a lot of the ordinary realism claims, and only retains realism about structure. None of my arguments are conclusive against Structural Realism, but they are not rejected. I find this similar to the position ordinary realism and constructive empiricism were held to be in. Both were philosophically acceptable positions regarding scientific theory, and there was no conclusive argument in favour of or against one of both.

I have my personal doubts about Structural Realism though. Firstly, it should be spelt out what structure is. As there are authors who have attempted this but my thesis is not about this subject, I will say no more than that. Secondly, I suspect whatever ‘structure’ means can be adjusted by defenders of Structural Realism in order to avoid Laudan’s argument. For example, does structure include the Lorentzian inertial frame of absolute rest? In order to avoid Laudan’s criticism it should not. But this decision to not include this in structure seems to me rather arbitrary. I suspect there is a self-fulfilling prophecy: Whatever fails to be retained as science develops, is not part of structure. Thus, structure is retained in the development of science, allowing us to be realists about structure.

However, I think such difficulties might be avoided with a good discussion about what structure is. In that sense I still have respect for the doctrine of Structural Realism and I will conclude with saying that it is probably a serious alternative to DONUD antirealism.

SOURCES

Books

Aristotle [Aristotle a], **Prior Analytics,**
http://classics.mit.edu/Aristotle/prior.1.i.html,
accessed on Wednesday 6/01/2011

Aristotle [Aristotle b], **Posterior Analytics,**
http://classics.mit.edu/Aristotle/posterior.2.ii.html,
accessed on 10/02/2011

Collins H. M., **Changing Order: Replication and Induction in Scientific Practice**
The University of Chicago Press, 1992, reprint from 1985 with a new afterword
ISBN 0-226-11376-0

Duhem P., **The Aim and Structure of Physical Theory**,
Princeton University Press, 1982, renewed from 1954
ISBN 0-691-02524-X

Earman J., **Bayes or Bust: a Critical Examination of Bayesian Confirmation Theory**
Massachusetts Institute of Technology, 1992
ISBN 0-262-05046-3

Feyerabend P., **Against Method**
Verso, 1993, third edition (original 1975)
ISBN 978-1-86091-646-8

van Fraassen B. C., **The Scientific Image**
Oxford University Press, 1980
ISBN 978-0-19-824427-1

Hume D., **Of Knowledge and Probability,** in *A Treatise of Human Nature*, Book I, Part III
http://www.gutenberg.org/files/4705/4705-h/4705-h.htm#2H_PART3,
accessed on 27/01/2011

Kant I., **The Critique of Pure Reason**
http://www.gutenberg.org/cache/epub/4280/pg4280.html,
accessed on 8-02-2012

Articles

Douven, I., "**Abduction**"
The Stanford Encyclopedia of Philosophy (Spring 2011 Edition), Edward N. Zalta (ed.),
URL = http://plato.stanford.edu/archives/spr2011/entries/abduction/

Fine A., **Unnatural Attitudes: Realist and Instrumentalist Attachments to Science**
in *Mind*, New Series, Vol. 95, No. 378 (Apr., 1986), pp. 149-179
Published by: Oxford University Press on behalf of the Mind Association
Stable URL: http://www.jstor.org/stable/2254025, Accessed on: 04/02/2011

van Fraassen B.C., **Empiricism in the Philosophy of Science**
in *Images of Science: Essays on Realism and Empiricism with a reply from Bas C. van Fraassen*, P.M. Churchland and C.A. Hooker (eds.), pp. 245-308, The University of Chicago Press, 1985

Gillies D., **The Duhem Thesis and the Quine Thesis**
in *Philosophy of Science: the Central Issues*, M. Curd and J.A. Cover eds, first edition, pp. 302-319
W.W. Norton & Company, Inc. 1998

Hacking I., **Do We See through a Microscope?**
in *Images of Science: Essays on Realism and Empiricism with a reply from Bas C. van Fraassen*, P.M. Churchland and C.A. Hooker (eds.), pp. 132-152, The University of Chicago Press, 1985

Hesse M., **Review: Changing Concepts and Stable Order**
Reviewed work(s): Changing Order: Replication and Induction in Scientific Practice by H. M. Collins
in *Social Studies of Science* Vol. 16, No. 4 (Nov., 1986), pp. 714-726, Published by: Sage Publications, Ltd., 1986, Article Stable URL: http://www.jstor.org/stable/285060

Kuhn T.S., **Objectivity, Value Judgment, and Theory Choice**
in *Philosophy of Science: the Central Issues*, M. Curd and J.A. Cover eds, first edition, pp. 102-118
W.W. Norton & Company, Inc. 1998

Kukla A., **Non-empirical theoretical virtues and the argument from underdetermination**
in *Erkenntnis* 41, 1994, pp. 157-1170
Published by: Kluwer Academic Publishers.

Kukla A., **Does every theory have empirically equivalent rivals?**
in *Erkenntnis* 44, 1996, pp. 137-166
Published by: Kluwer Academic Publishers.

Laudan L. and Leplin J., **Empirical Equivalence and Underdetermination**
in *The Journal of Philosophy*, vol. 88, No. 9 (Sep., 1991), pp. 449-472
Published by: Journal of Philosophy, Inc, 1991, Stable URL: http://www.jstor.org/stable/2026601

Laudan L., **Grünbaum on "The Duhemian Argument"**
in *Philosophy of Science*, Vol. 32, No. 3 / 4 (Jul. – Oct., 1965), pp. 295-299
Published by: The University of Chicago Press, on behalf of the Philosophy of Science Association
Stable URL: http://www.jstor.org/stable/186524

Laudan L. (a), **Demystifying Underdetermination**
in *Philosophy of Science: the Central Issues*, M. Curd and J.A. Cover eds, first edition, pp. 320-353
W.W. Norton & Company, Inc. 1998

Laudan L. (b), **A Confutation of Convergent Realism**
in *Philosophy of Science: the Central Issues*, M. Curd and J.A. Cover eds, first edition, pp. 1114-1135
W.W. Norton & Company, Inc. 1998

Lipton P., **Induction**
in *Philosophy of Science: the Central Issues*, M. Curd and J.A. Cover eds, first edition, pp. 412-425
W.W. Norton & Company, Inc. 1998

Maxwell G., **The Ontological Status of Theoretical Entities**
in *Philosophy of Science: the Central Issues*, M. Curd and J.A. Cover eds, first edition, pp. 1052-1063
W.W. Norton & Company, Inc. 1998

Muller F.A., **Can a Constructive Empiricist Adopt the Concept of Observability?**
in *Philosophy of Science*, 71 (October 2004), pp. 637-654

Muller F.A., **The Deep Black Sea: Observability and Modality Afloat**
in *The British Journal for the Philosophy of Science*, 56 (2005), pp. 61-99

Muller F.A. and Fraassen B.C. van, **How to talk about unobservables**
in *Analysis*, 68.3, July 2008, pp. 197-205

Musgrave A., **Realism Versus Constructive Empiricism**
in *Images of Science: Essays on Realism and Empiricism with a reply from Bas C. van Fraassen*, P.M. Churchland and C.A. Hooker (eds.), pp. 197-221, The University of Chicago Press, 1985

Norton J. D., **The determination of theory by evidence: the case for quantum discontinuity, 1900-1915**
in *Synthese 97*, pp. 1-31, Kluwer Academic Publishers, 1993

Norton J. D., **Must Evidence Underdetermine Theory?**
in *The Challenge of the Social and the Pressure of Practice: Science and Values Revisited*, M. Carrier, D. Howard and J. Kourany, eds., pp. 17-44, Pittsburgh: University of Pittsburgh Press, 2008

Norton J.D., **A Material Solution to the Problem of Induction**
From: http://www.pitt.edu/~jdnorton/homepage/cv.html 2009, accessed 16/05/2011

Quine W.V.O., **Two Dogmas of Empiricism**
in *Philosophy of Science: the Central Issues*, M. Curd and J.A. Cover eds, first edition, pp. 280-301
W.W. Norton & Company, Inc. 1998

Stanford P.K., **Underdetermination of Scientific Theory**
in *The Stanford Encyclopedia of Philosophy* (Winter 2009 Edition), Edward N. Zalta (ed.)
URL = http://plato.stanford.edu/archives/win2009/entries/scientific-underdetermination/
accessed on 11-04-2011

Stanford P.K., **Refusing the Devil's Bargain: What Kind of Underdetermination Should We Take Seriously?**
in *Philosophy of Science*, Vol. 68, No. 3, *Supplement: Proceedings of the 2000 Biennial Meeting of the Philosophy of Science Association. Part I: Contributed Papers* (Sep., 2001), pp. S1-S12
Published by: The University of Chicago Press on behalf of the Philosophy of Science Association
Stable URL: http://www.jstor.org/stable/3080930

Vickers J., **The Problem of Induction**
in *The Stanford Encyclopedia of Philosophy* (Fall 2011 Edition), Edward N. Zalta (ed.)
URL = http://plato.stanford.edu/archives/fall2011/entries/induction-problem/

Secondary Sources

Ladyman J., **What's Really Wrong with Constructive Empiricism? Van Fraassen and the Metaphysics of Modality**
in *The British Journal for the Philosophy of Science*, 51 (2000), pp. 837-856

Kuhn T.S., **The Structure of Scientific Revolutions**
The University of Chicago Press, third edition, 1996
ISBN 0-226-45808-3

Feyerabend P.K., **Conquest of Abundance: a tale of abstraction versus the richness of being**
The University of Chicago Press, 2001
ISBN 0-226-24534-9

Shapin S. and Schaffer S., **Leviathan and the air-pump: Hobbes, Boyle and the experimental life**
Princeton University Press, 1989
ISBN 0-691-02432-4